Visually Guided Control of Movement

NASA

Visually Guided Control of Movement

Edited by
Walter W. Johnson
Mary K. Kaiser
Ames Research Center
Moffett Field, California

NASA

CONTENTS

PRECEDING PAGE BLANK NOT FILMED

CONTENTS (concluded)

ABSTRACT

The papers in this volume were presented at an intensive, three-week workshop on visually guided control of movement. The participants were researchers from academia, industry, and government, with backgrounds in visual perception, control theory, and rotorcraft operations. The papers include invited lectures and preliminary reports of research initiated during the workshop. Three major topics are addressed: extraction of environmental structure from motion; perception and control of self motion; and spatial orientation. Each topic is considered from both theoretical and applied perspectives. Implications for control and display design are suggested.

AGENDA

June 26

9:30	Walter Johnson	Welcome and introduction of workshop personnel
10:00	Sandy Hart	Rotorcraft Human Factors Program Overview
10:45	James Vorhees	Visual Cues used in Rotorcraft Flight
11:30	Lunch	Galileo Room in Cafeteria
1:30	Ron Hess	Feedback Models of Human Pilot Control Behavior
2:30	Rik Warren	Optical Information for Egomotion
3:30	James Cutting	Wayfinding from Optic Flow
6:00	Workshop BBQ by pool at St Francis Arms Apartments (dinner at 7:00)	

June 27

9:00	Ian Howard	Determinants of Spatial Orientation
9:45	Lawrence Hettinger	Applied Aspects of the Study of Spatial Orientation
10:30		Organizational Meeting

Afternoon — Initial Group Meetings

June 28

9:00	Steve Ellis	Perception of Egocentric Direction from Maps
9:45	Vernol Battiste	Geographical Orientation
10:30		Tour of laboratory facilities
11:30	Lunch	Galileo Room in Cafeteria

Afternoon — Groups I and II Joint Meeting/Individual Efforts

June 29

9:00	Dean Owen	Perception and Control of Self Motion
9:45	Tom Bennett	Visual versus Optical Flow: Where is the Real Ambient Array?
10:45	Groups II and III Joint Meeting/Individual Efforts	

June 30

9:00	John Andersen	
9:45	Lawrence Hettinger	Optical Variables Underlying Vection
10:45	Groups I and III Joint Meeting/Individual Efforts	

June 27 – July 12

During these days the workshop will meet as a whole at 8:30 AM and will hear one to two 30-40 minute talks by participants. Following this group meeting, individual meetings and individual work will take place.

July 12 – July 14

Final reports by participants to the workshop as a whole.

Evening July 13 — Workshop BBQ by pool at St Francis Arms Apartments

WORKING GROUPS

Group I – Information about 3-D structure from passively and actively sampled optic flow

John Perrone

Mary Kaiser

John Andersen

Dennis Proffitt

Joe Lappin

Group II – Control of locomotion through direct use of optical relationships or through use of optically derived 3-D representations

Walter Johnson

Tom Bennett

Ron Hess

John Flach

Lawrence Wolpert

Dean Owen

James Cutting

Group III – Sources and determinants of spatial (dis)orientation

Steve Ellis

Rik Warren

Ian Howard

Greg Zacharias

Gary Riccio

Lawrence Hettinger

INTRODUCTION

Understanding how people use visual information, either sampled through windows or from pictorial displays, to control vehicular motion is critical to the way helicopters are designed for takeoff, landing, and low-level flight. During the three weeks from June 26 to July 14, 1989, the Rotorcraft Human Factors Research Branch at the NASA Ames Research Center in Mountain View, California, assembled a talented cross section of researchers from academia, industry, and government to examine, develop, and prepare tests of fundamental ideas in this area.

The Summer 1989 Workshop on the Visually Guided Control of Movement was unusual in both its length and in the intensity of the interactions among its participants. Traditional workshops tend to last no more than 4 or 5 days. The main problem with such meetings is that there is only sufficient time for participants to state their positions or theories, or to report on their latest experiments. They do not allow much in-depth discussion of issues, especially when the problems are difficult and/or the participants have different orientations. Often the participants go away without having had the opportunity (or the challenge) to resolve exactly how their viewpoints differ. Instead, they are left wondering if they have missed a vital point, or if a fellow participant is badly misguided.

We decided to try something different. It seemed to us that three weeks should be enough time for all participants' ideas and positions to be sufficiently clarified. But we knew that just providing time would not be enough. A structure was needed that would focus interactions but would not inhibit the free flow of information and thought. Therefore we organized our workshop about two poles. First, each day, the participants met as a body, and one or two participants gave an informal presentation. Second, there were daily meetings of three groups formed to explore three major research topics: (1) the perception of structure from motion, (2) the perception and control of self-motion, and (3) determinants of spatial orientation. Furthermore, since all participants were staying at the same apartments, many valuable discussions occurred outside scheduled workshop hours.

Since none of us had heard of a similar undertaking, we knew we were taking a risk. While we hoped that the extended length could counter the shortcomings of traditional conferences and workshops, many participants and organizers were not fully confident that intellectual momentum could be maintained beyond a single week. Judging from the comments of the participants, however, the workshop exceeded the expectations of all involved. By the end of the third week all agreed that they were leaving the workshop with momentum intact.

The papers included in this publication are not the final results of the workshop. They are, instead, samples of the issues discussed during the workshop. They include theoretical work as well as proposed experiments. During the workshop, the participants were strongly encouraged to generate research designs. This mechanism was used to focus people with diverse theoretical interests upon a common topic, thereby making communication necessary.

First, Mary Kaiser gives a personalized summary of the workshop.

The contributions of Hart and Battiste, Ellis, and Bennett examine issues relevant to helicopter navigation and flight. The report by Hart and Battiste is a continuation of one of the more popular

workshop presentations about the skills used during low-level helicopter flight and navigation. Their report helps embed and contextualize the more abstract or basic issues into the applied content area of helicopter flight. Ellis reports on more basic experiments in which he used map-like pictures to examine how the relative positions of objects were influenced by map geometry. The report by Bennett is a further effort in the more applied vein, dissecting the pilot's task into psychologically important component tasks and skills.

The issue of extracting structure from motion is explored in the related reports of Lappin and Perrone. Lappin proposes a model of how humans extract 3-D structural information, and information about their own self-movement, during self-movement. John Perrone provides an analysis of the visual perception of surface slant during self-movement.

The reports by Cutting and Owen are continuations of workshop discussions about whether human visual perception is anchored to retinal or optical arrays or frames of reference. This was one of the central debates of the workshop.

The reports by Andersen, Wolpert, Johnson and Phatak, Hess, Zacharias, and Flach all reflect concerns with more explicit descriptions of the visually guided active control of movement. Andersen provides a discussion of the relevant visual information for the support of various flight activities. Wolpert discusses the relative utility of multiple sources of visual information for the control of altitude during forward flight. Hess examines the control of lateral heading in automobile driving and extends his findings to helicopter control. Johnson and Phatak, Zacharias, and Flach all delve more deeply into the nature of the control models underlying the control of self-movement (Anil Phatak was not able to attend the workshop, but was instrumental in laying the groundwork for it).

Proffitt's report is concerned with the importance of context in phenomenal perceptual experience and how this relates to various issues. One of these issues, which was a subject of workshop discussions, is the perceptual penetrability of physical dynamics. His previous work in collaboration with Kaiser suggests that the perception of physical dynamics (e.g., energy, momentum, mass) is highly constrained. This topic formed another central debate of the workshop and is discussed in several other reports, in particular those of Owen and Flach.

Riccio, Hettinger, and Howard all discuss issues related to spatial orientation. Riccio argues for the importance of nonvisual sources of information in a pilot's maintenance of spatial orientation. Hettinger discusses the relationships among vection, interperceptual correlations, disorientation, and motion sickness. Howard provides a discussion of the functions of egocentric and exocentric frames of reference in the maintenance of spatial orientation.

REFLECTIONS ON THE WORKSHOP
(A Personalized Summary)

Mary K. Kaiser
NASA Ames Research Center

The 1989 Workshop on the Visually Guided Control of Movement provided a forum for researchers with different ways of thinking and talking about similar problems to interact, react, and (perhaps) rethink. It provided many of us the opportunity to gain a cursory education about new areas, such as control theory, and to understand more about the rotorcraft environment and typical pilotage tasks. In order to focus participants' discussions, Walt Johnson and I proposed that participants consider five questions that the helicopter pilot must solve:

1. Am I going?

2. Where am I going?

3. How fast am I going?

4. What environment am I going through?

5. Which way is up?

These questions, though seemingly simple, actually capture the most important issues of the workshop. "Am I going?" addresses the topic of vection. Ian Howard presented some preliminary data which indicate that optikinetic nystagmus (OKN) is decoupled from vection, and that it is driven by different aspects of the visual display. John Andersen has done some intriguing work on people's postural adjustments to vection. He and Ian agreed that surfaces perceived as most distant determine vection response, and that other factors previously implicated (e.g., visual angle of display) are not as critical as once thought.

"Where am I going?" addressed the issue of wayfinding and extraction of heading. James Cutting presented a model of terrestrial wayfinding, which several participants thought could be extended to helicopter navigation tasks. There has been much discussion (some of it rather heated) about how people extract heading from optical information. Plans are afoot to design some empirical studies comparing the competing models, which should generate light instead of heat.

The question "How fast am I going?" got somewhat confusing, because in some of the work presented, subjects controlled altitude as well as (or instead of) velocity. Walt Johnson reported that subjects used edge rate rather than flow rate, even when edges were stochastically rather than regularly placed. He agreed that flow rate can have an effect, but it seemed to be overshadowed by edge rate in his displays. As with heading extraction, further studies are being planned.

"What environment am I going through?" was the question asked by those of us concerned with the problem of extracting structure from motion. I was joined by John Andersen, Joe Lappin, John

Perrone, and Denny Proffitt. After several days of preliminary discussions, we decided upon a study of mutual interest, which I will describe later. At first, we talked about a number of topics of interest to us. For example, Denny has become very interested in the stereokinetic effect (SKE), a phenomenon that has been discussed in the literature for a long time (several Italian researchers studied it in the 1920s), but never understood very well. He is interested in the SKE because people have very compelling and stable form percepts from visual motion stimuli which do not map to any realizable rigid geometry (and certainly not to the perceived geometry). Hearing him, Joe, and the two Johns discuss their geometric analyses of motion displays drove home the Putnam quote James Cutting used at his discussion of direct vs. directed perception: there really are multiple descriptions that must be considered. Anyway, we eventually were able to find a topic of interest which transcended paradigmatic differences: slant perception.

"Which way is up?" really concerns more than up, and, of course, in the navigated platform context there can be more than one up. Ian Howard has done (and continues to do) very interesting work on the problem of orientation. Gary Riccio presented a study which had a very clever decoupling of vestibular gravitational cues and inertial balance cues. Also, Irv Rock came down from UC Berkeley one day, and that added a spirited discussion.

An ongoing topic of debate was whether perception examined in the context of passive observation was equivalent to perception in an active performance context. This debate brought to light several important issues concerning the use (and lack thereof) of active control measurements in research.

The argument for active control measurements can take several forms. Perhaps the strongest version is that people utilize *different* information in making passive, verbal judgments than when asked to perform some action in context. Gary Riccio would perhaps make such an argument—his subjects reported that they were gravitational tilted, yet the control data suggest that they could maintain a gravitationally upright orientation. Dean Owen seemed to support this strong argument as well, and contended that he no longer studied "responses" to stimuli. This, of course, caused much consternation among those of us who employ passive observation paradigms.

The rebuttals took several forms. James Cutting felt that there was very little empirical support for the idea of differential information usage in active and passive contexts, and considered the argument that such differences exist to be a "promissory note" at this juncture. Denny Proffitt expressed a more fundamental concern that control models do not adequately characterize the task being studied, and they reduce most tasks to the level of error correction. He feared that this was no more than Hullian stimulus–response modeling, with subjects performing corrections to unanticipated disturbances. Such a limitation to the sophistication of control modeling would seriously compromise its ability to fulfill the promise of delineating which information subjects are using to control behavior.

I, too, have several reservations regarding the supposed superiority of active control experimentation. It seems to me that the experimenter still must select the information potentially available in the display. If, as Dean suggested, this selection is based on verbal responses of pilots, little is gained with regard to distancing the research from the phenomenology. Second, I agree with James Cutting that the time course analyses fundamental to control models are made problematic by motion

information which lacks clear-cut onset times. Finally, it seems to me that active control research examines the performance that is, instead of the performance that could be, and both kinds of performance are of interest to the psychologist. In an active control context, subjects may employ highly suboptimal strategies for sampling of the stimulus space. I guess I think John Andersen's voice of reason is pretty appealing ...active and passive experimentation can provide converging sources of information, but neither has any intrinsic privilege.

Another interesting difference among the participants was how they characterized what subjects know about what they are doing. Perhaps this difference comes from the fact that some of us work with highly skilled subjects, e.g., pilots, who are very good at articulating what they are doing (or at least what they *think* they are doing) when performing a task, while others work with naive populations who are less articulate. This brought up some interesting considerations of the extent to which the dynamics of a system must be accessible to someone utilizing it. Denny Proffitt argued that there is no evidence or logical requirement that a person needs an adequate representation of system dynamics; one need only have a transform function which successfully maps actions to the resulting kinematics. (Of course, as he himself admitted, a similar argument can be raised about perception: as long as a person maps the perceived situation onto appropriate behavior, it matters little whether or not the perception is "veridical" in any formal sense of the word. A possible example of this phenomenon seems to occur in distance perception, where a person may verbally underestimate distances, yet successfully walk the correct distance to a point when blindfolded.) This issue was not resolved; the control people still have their operator model boxes filled with explicit formulations of system dynamics while Denny maintains that the box is filled with heuristics and action/kinematics mappings.

Since the control theory could not persuade us passive structure-from-motion types of the error of our ways, we designed a study on slant perception which does not utilize active control (at least in the initial paradigm), but is a fairly interesting study anyway. Denny Proffitt actually brought up the topic, although John Perrone has done a good deal of work on the topic (and has a related study ongoing with Walt Johnson). Denny was thinking of exemplars of everyday misperceptions which might have important implications for rotorcraft navigation. He suggested that people have a tremendous tendency to overestimate the slant angles of hills (relative to horizontal). For instance, if you ask most people to estimate the steepest slope in San Francisco that they drive on, most will respond with a figure far greater than the actual value of 15 degrees. This does not seem to be an artifact of memory, because similar overestimates are given when people are actually looking at a hill.

At a sufficient altitude the opposite effect may occur: from an altitude of 30,000 feet, mountains appear almost flat. Thus, we decided that approach altitude would be an important variable to consider. Next, we talked about what information actually specifies slant. John Perrone had considered linear perspective information, and was now getting interested in motion-based information. After several days of discussion, we determined that relative slope could be recovered from motion, and set about programming displays that would vary the slope of a ramp relative to a horizontal ground plane.

The plane and the ramp are defined by point-lights randomly distributed on their surfaces. This means that there is texture gradient information concerning the slope, thus a static control is required. John has altered texture gradient cues in his other display, but we wanted to keep texture

information naturalistic and evaluate its utility via the static control. Both our subjective impressions and the literature on slant perception suggest that texture density *per se* is not sufficient to specify slant. We have three altitudes of approach, simulating the eye heights of low-, medium-, and high-altitude rotorcraft flight.

I, myself, have some difficulty perceiving the slant when approaching the ramp head-on (i.e., in a z-axis approach). I do not always perceive the ramp as rigid, and such rigidity is required to extract slant. Furthermore, the texture gradient seems rather uninformative (not surprisingly), and sometime even seems to work against the depicted slant. My faulty perception may have led, in part, to the third factor we have included: translation axis. Thus, we have x-axis as well as z-axis translations. The x-axis translation produces parallax information which seems much more informative about slant (and greatly mitigates the overestimation bias). John Perrone, in addition to having pro-grammed the study, is performing motion analyses on the two translations. He suspects that much of the slant-specifying motion in the z-axis translation is subthreshold (and even subpixel, on our display system).

In fairness, we have thought about the issue of active control. First, Denny (perhaps at my prod-ding) brought up the issue of whether such slant misperception would have any consequences on performance. After all, when we approach the hill in San Francisco, we do not lift our legs too high and fall on our faces. Similarly, the helicopter pilot may see a 15-degree slope as 40 degrees, but as long as his control input is appropriate for the 15-degree slope (and it will be if he has learned the proper control response for a hill that looks like the one he is currently viewing), the calibration of his percept relative to the objective geometry is irrelevant. I guess we sort of finessed this question by agreeing that there are many instances in which one would like to acquire an accurate impression of terrain layout, so it is relevant to determine how these impressions are affected by approach alti-tude and direction. We will probably deal with control issues in later studies, but first we want to document the basic phenomenon. If these factors do affect veridicality of slant perception, I would be interested to see whether people adopt an optimal sampling strategy when left to their own active devices.

So we have a four-factor, within-subject design. Displays will either be static (showing the mid-dle frame of the trajectory) or contain motion. Translations can occur along the x or z axis (later studies will utilize oblique translations) at three altitudes. Eight levels of slant will be used (15–120 degrees in 15-degree intervals). We decided to use angles greater than 90 degrees in order to access whether the slant continues to be overestimated or is, rather, biased toward vertical. Observers will respond by setting the slope of a ramp depicted orthogonal to the display's orientation.

John Perrone and I will collect the data after the workshop ends. We will have preliminary analy-ses done by the time we go to the Psychonomic Society meeting in November, and decide how to proceed at that point. Chances are we will want to pursue different issues based on these preliminary findings. It is really great, though, that we got the opportunity to start the project at this workshop. I think we have all learned a lot from the experience, and have even managed to plan some good science. As an organizer of this workshop, I could not have asked for more.

THE USE OF VISUAL CUES FOR VEHICLE CONTROL AND NAVIGATION

Sandra G. Hart and Vernol Battiste
NASA Ames Research Center
Moffett Field, California

INTRODUCTION

At least three levels of control are required to operate most vehicles: (1) Inner-loop control to counteract the momentary effects of disturbances on vehicle position, (2) Intermittent maneuvers to avoid obstacles, and (3) Outer-loop control to maintain a planned route. Operators monitor dynamic optical relationships in their immediate surround to estimate momentary changes in forward, lateral, and vertical position, rates of change in speed and direction of motion, and distance from obstacles. They seek, identify, and locate specific landmarks to maintain more global geographical orientation. Mental rotation and transformation may be required to align information in maps, instruments, or memory into alignment with the visible scene for comparison. The process of searching the external scene to fmd landmarks (for navigation) is intermittent and deliberate, while monitoring and responding to subtle changes in the visual scene (for vehicle control) is relatively continuous and "automatic." However, since operators may perform both tasks simultaneously, the dynamic optical cues available for vehicle control task may be determined by the operator's direction of gaze for wayfmding.

Constraints imposed by the mission, the vehicle, and the environment determine the temporal and spatial precision with which operators can and should execute their activities, the information that is available, and the processes by which navigation and immediate control are accomplished. Routes may be explicit and visible in the external scene (i.e., roads), represented on displays in digital or analog formats (i.e., air routes), or evolve in response to information obtained and events that occur during the mission (i.e., maneuvering around unexpected obstacles) . Operators rely on a variety of information sources and reference systems to accomplish each level of control. However, the utility of information for different control functions varies within and between missions, depending on the operator's goals and experience and the unique characteristics of the vehicle and the environment.

The following is an attempt to relate the visual processes involved in vehicle control and wayfinding. The frames of reference and information used by different operators (e.g., automobile drivers, airlme pilots, and helicopter pilots) will be reviewed with particular emphasis on the special problems encountered by helicopter pilots flying nap of the earth (NOE). The goal of this overview is to describe the context within which different vehicle control tasks are performed and to suggest ways in which the use of visual cues for geographical orientation might influence visually guided control activities.

AUTOMOBILE DRIVERS

When driving a car, the current route and choice points are immediately visible. Furthermore, target performance criteria are well defined: (1) Speed limits are posted or drivers may match their speed to the flow of traffic, and (2) Lateral position is constrained by the width of the road or the driver's lane.

Navigation

To maintain geographical orientation, an automobile driver's knowledge of an area does not have to extend very far beyond the road system. If he is on the correct road, traveling in the correct direction, and can recognize relevant choice points, he does not need to know exactly where he is most of the time nor anything about the streets, structures, or terrain features on either side of his route. Drivers need to refer to other coordinate systems (e.g., compass direction) only when making decisions about which way to turn at an unfamiliar intersection where options are distinguished by North/South or East/West. In most cases, drivers can navigate well even at night, in poor visibility, and in unfamiliar areas because their options are limited by the structure of the road system.

Thus, the mental models drivers develop of their environment are composed of major arteries (e.g., their names, orientation, or end points), the relationships among them (e.g., significant intersections, or relative orientations and distances), and detailed information about secondary roads in specific areas. They may organize information about isolated groups of familiar secondary roads by their proximity to major arteries, specific places, or geographical features. In addition, people can infer the location of an unfamiliar place if streets are laid out in a regular pattern and named in a logical sequence. Automobile drivers develop mental models of familiar environments through experience. They elaborate these models over time, incorporating new information about previously unfamiliar areas or additional information about familiar areas.

When driving from one place to another, people plan and follow a route by referring to: (1) remembered or written route lists (e.g., street names, turn directions, and time or distances between turns; (2) remembered spatial relationships among streets (whose names may not be known), (3) visible landmarks, and/or (4) maps. When driving to an unfamiliar place in a generally familiar area, they can develop an approximate route based on their general knowledge of the area, while they must rely on explicit instructions or a map in an unfamiliar area.

Figure 1 depicts a typical road map used by automobile drivers. Figure 2 depicts a more spatially compatible perspective view that integrates major highways with significant terrain features and landmarks. The latter type of map provides a driver, that is unfamiliar with an area, with explicit cues about how landmarks will look and the relationships among traffic routes, terrain, and significant cultural features.

Automobile drivers are generally free to choose any route they wish and deviate from a planned route at any time; there are no externally imposed constraints on departure times, route selections, or route changes. Enroute, they may verify that they are on course by identifying features along the way or reading road signs. If they are not sure where they are, they may have sufficient general

knowledge of the area to locate a familiar feature to re-orient themselves. The selection (or change) of routes and departure, enroute, or arrival times are usually based on personal time constraints (e.g., a desire to arrive at work on time). To maintain a schedule, drivers estimate where they are, the distance from their destination, and probable driving time based on past experience or mental arithmetic. If they encounter traffic congestion or road construction they may switch to another route or adjust their speed. The number of options available to drivers are determined by the availability of alternate routes and their knowledge of the area.

In an unfamiliar area, drivers may use cues and representations that are similar to those used in familiar areas, but their knowledge of the environment is limited to a few highways, significant intersections, and landmarks. Their mental models are sparse and may be based solely on the quick review of a map. Their time/distance judgments are likely to be less accurate and they have limited flexibility if they encounter problems using the planned route. If they miss a turn, or turn in the wrong direction, they may have to retrace their steps or consult a map to figure out where they are.

When giving directions or acting as a navigator from the passenger's seat, people generally refer to roads or places by name and give instructions oriented to the driver's frame of reference ("Turn right at the stop sign."). They may refer to compass directions to improve the general geographical orientation of the recipient ("The park is 2 miles South of the intersection.") or identify a specific location ("The store is located on the Northeast side of the intersection."). They may provide additional information about boundaries ("If you pass the mall, you have gone too far."), choice points ("Turn right on 1st Street just past the park."), or distances ("The intersection is in 2 miles."). Finally, they may provide predictive information to allow the driver to plan ahead ("Stay to the right after the bridge."). In most cases, people use explicit names and distinctive, visible features to aid recognition. This process is facilitated if both individuals share a common knowledge of the area. If they do not, then verbal labels may have to be supplemented with a description of significant landmarks.

Vehicle Control

In an automobile, drivers rely on visual cues for both vehicle control and navigation, rather than on instruments. They continuously scan the environment to avoid obstacles and regulate speed and lateral position. Although they can refer to the speedometer to determine their actual speed, most control inputs reflect estimates of absolute speed, relative speed (in comparison to other automobiles), or changes in speed that have already occurred or will occur (e.g., when approaching hills or slower traffic). These estimates may be based on optical cues (e.g., optical flow, edge rate, rate of closure with moving or stationary objects), auditory cues, or vibration. The accuracy of such estimates may be reduced when operating an unfamiliar automobile; if the driver's eye height is significantly higher or lower than usual (because the vehicle is a different size), there may be a consistent bias in speed estimates.

Lateral control is primarily based on optical cues; drivers generally try to remain centered in their lane and safely separated from other traffic. When driving in a cross wind, drivers compensate by adopting a constant bias in their control input. The frequency with which lateral control inputs are required depends on the road surface and traffic density. Required control precision depends upon lane width, car width, and traffic density.

AIRLINE PILOTS

The pilots of commercial jets are faced with a different situation. They fly high above the earth where there are no visible routes to follow and environmental cues are few and far between. Although they could use the sun and stars for navigation, celestial navigation is difficult, imprecise, and impossible when the sky is obscured by clouds. Alternatively, they might refer to significant landforms to improve their geographical orientation. However, these cues might be too distant to use as a primary cue or invisible in poor weather or at high altitudes. Thus, pilots generally rely on instruments for navigation and flightpath control.

Navigation

Given the increasing density of air traffic, greater navigational precision and coordination have become necessary. Thus, formal route structures have been created that are defined by arbitrary coordinate systems referenced to agreed upon standards (e.g., magnetic north) and a network of navigation aids. Information from these sources provide "pathways" for pilots to follow which are not directly visible but instantiated on instruments, displays, and charts.

Pilots must integrate dynamic information presented in different formats (digital/analog), spatial dimensions (one-dimensional/two-dimensional), and units (knots, degrees, feet) that are referenced to many different coordinate systems (earth referenced—intertial, magnetic or polar coordinates; vehicle referenced—longitudinal, vertical, and lateral axes) to develop a dynamic, three-dimensional mental model of the environment. Furthermore, traditional cockpit instruments are not referenced to the ground below. Thus, pilots must infer their position and ground speed. For example, a magnetic compass displays heading rather than ground track; winds may cause the craft to drift off course, while the aircraft's heading remains constant. Airspeed indicators display rate of movement through the air rather than across the ground. Barometric altimeters display height above sea level rather than height above landforms immediately below the aircraft.

In general, airline pilots' knowledge about their location is referenced to these (invisible) route structures, which are superimposed upon, but not necessarily related to terrain features. These systems allow very precise navigation, even when visibility is zero, but require the human operators to maintain very complex mental models of their environment. Because the air route structure is the basic reference system, rather than visible terrain features, pilots may not be "lost" even if they have no idea what state they are flying over; as long as they are on time and on course, they know all they need to know. As with automobile drivers, pilots' mental models of the environment, and the degree of precision with which they must maintain geographical orientation is substantially constrained by the route structure within which they operate. However, they may also incorporate information about terrain features, weather systems, and other vehicles (from visual observation or radio communications) into their mental models.

In aviation, flight plans are based not only on altitudes and bearings, but also on time. In order for the air traffic control system to operate smoothly, pilots must depart and land on time, and arrive at "fixes" (imaginary points in the sky that represent the intersection of two radio navigation signals) on schedule. Although these nominal times are worked out in advance, based on the aircraft's speed

and predicted wind conditions, the situation may change. Thus, pilots may have to adjust their speed to stay on schedule. However, in conventional aircraft, pilots must infer the distance they have traveled across the ground, as their instruments display airspeed, rather than groundspeed.

Enroute, pilots communicate about their current position and planned route within the context of these arbitrary reference systems (e.g., heading, distance from a navigation aid, arrival at a "fix"). The language they use is highly structured and constrained to facilitate accurate and rapid transmission of information. They maintain geographical orientation by correlating the information viewed on their instruments with paper charts. Figure 3 depicts a high-altitude chart used in flight above 18,000 ft.

The only time pilots must adopt a frame of reference based on directly visible cues is during landing. At this point, they must transition from one mental model (based on an arbitrary route structure) to another (visible structures and terrain features viewed in the external scene). In addition, they may compare visible cues to those depicted on an approach plate. Figure 4 depicts an approach plate used when landing at an airport. It includes some information about visible landmarks as well as the route the pilot is to follow. After transitioning to a visual frame of reference, pilots may report their position with respect to visible landmarks whose location is likely to be known by the message recipient.

Vehicle Control

During high-altitude flight phases, airline pilots base their manual control inputs on dynamic optical cues displayed on instruments; speed, altitude, and course are regulated by detecting and reducing errors between the target value and the current value. In some cases, the same instruments are used for vehicle control as for navigation. The effects of wind on ground speed and ground track must be inferred.

The spatial relationship between movement of an indicator on an instruments, control inputs, and movement through space are often incompatible. For example, the effects of right/left control inputs to changes in heading are reflected in rotation of the compass in the opposite direction (the display is "inside-out"). Fore/aft throttle inputs are reflected in rotations of the airspeed indicator (clockwise, faster; counterclockwise, slower). Fore/aft inputs in the control yoke and/or the throttle affect attitude and power, which determine altitude. The altimeter depicts height above the ground (radar altimeter) or above sea level (barometric altimeter) in two formats: digital readout (coarse-grained) and circular dial (fine-grained). Flight-directors are the only instrument that provides information about pitch, roll, yaw, and deviation from desired course in a spatially compatible format. However, these displays are "inside-out" (e.g., the "world" moves, while the "aircraft" remains stationary in the center of the display) and two dimensional, rather than perspective.

Although each instrument provides information about a specific dimension (e.g., altitude, airspeed, attitude), control inputs may influence more than one dimension (e.g., changes in altitude will also affect speed unless pilots compensate by adjusting the power setting). Rather than entering constant adjustments, most pilots wait until error has exceeded a criterion value; in most cases, smooth control is more important than precise control, to ensure passenger comfort. In all modern aircraft, autopilots allow pilots to set desired values by entering discrete commands; automatic subsystems

achieve and then maintain the selected values at the specified times. Pilots simply monitor the system to ensure that it is functioning properly. When acting as either manual controllers or monitors of automatic systems, pilots must maintain an integrated, multidimensional model of vehicle state based on input from many sources expressed in different units of measurement and reference systems.

Only in the initial and final phases of flight, when departing from or approaching an airport, do pilots refer to dynamic optical cues visible in the external scene to monitor lateral position, altitude, and speed. Their task is more complex than that of automobile drivers: (1) They must worry about additional degrees of freedom (e.g., height above the ground, attitude, bank angle); (2) They are traveling three to four times faster and, thus, require greater visual range; and (3) They must relate their estimates of vehicle motion based on dynamic optical cues in the external scene to values displayed on instruments.

HELICOPTER PILOTS

The pilots of military or civilian helicopters flying at very low altitudes are faced with an even more difficult situation. They operate so close to the ground that local terrain features may obscure their view of significant landmarks and restrict their visual range. This makes it difficult to relate local terrain features to a more global context. Often, helicopters move freely through terrain, without an explicit (visible or electronic) route to follow. While there are many degrees of freedom in this environment (helicopter crews are not limited to roads or electronic routes), it is more difficult to maintain the desired course and natural and man-made obstacles pose a very real threat. In this environment, helicopter crews must correlate cues viewed in the external scene with information on paper maps to maintain geographical orientation, avoid obstacles, and maintain their course. Instruments that provide pilots with information about speed and altitude are relatively inaccurate at low altitudes and slow speeds and electronic aids must have a line of sight with the source to work properly.

Navigation

Before a mission, helicopter crews study maps of the environment in which they will operate to select a route that offers the most direct path to the destination (given terrain contours, obstacles, etc.), distinctive visual cues (to aid in geographical orientation), and cover (if there is an enemy threat). They select specific features that they will use during the mission to verify their location and identify choice points (e.g., intersections of rivers, hill tops, clearings, groves of trees). They might identify linear features that can provide a visible "route" to follow (e.g., ridge lines, river valleys). Military crews avoid selecting man-made structures for reference (things change) and following roads (the enemy threat is greater there).

Helicopter crews incorporate available information into a cognitive model or mental map of the environment through which they will travel. The mental representation might be spatial—a mental image of the map (a plan view) or a series of perspective mental images of how significant features in the environment are likely to look when viewed from the cockpit of a helicopter (a forward view). Alternatively, they may store this information as a route list—a series of verbal commands (e.g.,

"Travel down the valley for 2 miles then bear right") or descriptions (e.g., "Follow the creek that runs beside the cliff") that are remembered and executed during the mission.

During a mission, helicopter crews view features in the external scene and compare them to a paper map or their mental images. They must mentally transform the stylized images on two-dimensional maps into mental images, that represent a perspective view of the object. The image is then mentally rotated to bring it into alignment with the forward field of view for comparison with the external scene. If they continue to see expected features on time and in the correct order, they know where they are; visible terrain features correspond with their expectations and they can corre-late their position with a location on the map. For example, when they pass a distinctive feature (e.g., a water tank depicted on their map) or intersecting linear features (e.g., two ridge lines), they know precisely where they are. However, if a single landmark is symmetrical, they may know generally where they are, but not their precise location or the direction from which they are approaching the feature. In this case, they may look for a second reference point, check the compass, look at the sun, or infer direction from previous cues. When using a ridge line that extends for some distance as a geographical reference, a crew only knows that they are traveling in the correct direction, but not their precise location.

Depending on the familiarity of the terrain, the availability of distinctive features, and the quality of pre-mission planning, maintaining a route may be relatively easy or very difficult. For example, when a crew must rely on subtle variations in terrain to judge location, it may be extremely difficult to relate features visible in the forward scene to contour lines on the map. This task is particularly difficult if surface contours are masked by vegetation. Furthermore, the appearance of terrain and vegetation varies seasonally and from one region to another, requiring adaptation and inference. There may be considerable ambiguity about whether a particular feature is, in fact, the one a crew expects to see, or the specific feature depicted on the map.

As the time between landmarks increases, uncertainty about current position may increase if additional cues are not available for the crew to verify that they are, in fact, where they think they are. At some point, the crew will begin to look for the next expected landmark. If it does not appear by the expected time, the crew may begin to consider the possibility that they are lost. If a feature that is similar to their expectations appears, the crew may identify it as the expected feature. If it is not, it make take some time before they accept the growing evidence that they are not where they are supposed to be. At this point, the crew must take action to re-establish their position. A helicopter pilot might gain altitude to find a distinctive landmark. If this is not possible, he may carefully sur-vey the surrounding terrain and try to find a pattern of features on the map that corresponds to what he sees. However, it is much more difficult to find a pattern somewhere on a map that corresponds to the forward scene, than to verify that a visible feature is where it is supposed to be relative to the vehicle. Alternatively, tne pilot may try to re-trace his path until he finds a familiar landmark. How-ever, the mental preparation performed before the mission will be of little help here, as terrain features and relationships will not correspond to the expected sequence or orientation.

Thus, maintaining geographical orientation requires helicopter crews to continuously correlate the visual scene with the map. Estimates of when to begin looking for a landmark, whether a choice point has been missed, or what features should be visible at any point in time are based on subjective

estimates of the distance traveled and the time elapsed since the last known location. Explicit calculations are difficult because the route might not have been direct nor followed a straight line.

When operating at night, helicopter crews rely on night vision devices (that intensify light or display infrared imagery) to provide them with information about the external scene. Although they could not perform required missions without these devices, their use imposes considerable additional load on the pilots; field of view is limited, acuity is reduced, depth cues are distorted, subtle textures necessary to identify a particular feature may be missing, and objects or terrain features may look very different than expected. Furthermore, greater navigation precision is required at night; obstacles that can be seen and avoided during the day may be invisible at night. Thus, pilots rely on maps to spot potential obstacles. However, this information is useful only if they know exactly where they are. For these reasons, maintaining geographical orientation becomes significantly more difficult and overall prrformance capabilities may be reduced. For example, pilots are more likely to fly slower and higher at night.

In helicopters, crewmembers convey information about navigation and geographical orientation verbally, although they may use gestures, as well (e.g., point to features in the environment or on a map). In NOE flight, navigation may take as much as 90% of the navigator's time, and communications between the pilot and navigator about navigation, 25% of both crewmembers' time.

Army aviators use 1:50,000 scale maps (Figure 5) that depict terrain contours (e.g., hills, valleys), vegetation (e.g., fields, groves of trees) bodies of water (e.g., rivers, streams, ponds), and some cultural features (e.g., roads, buildings, bridges, water tanks, towers). During pre-mission planning, helicopter crews plot their route on the map, identify critical choice points, and select additional features that they will use to verify their position. In flight, the navigator follows the route of flight on the map, giving the pilot verbal cues about what he should see, when he should begin or end a turn, and potential obstacles. In addition, the navigator scans cockpit instruments, verbalizing relevant information to the pilot. The pilot generally keep his eyes on the forward scene, telling the navigator what he sees and verifying that he can (or can not) see a specific landmark.

Helicopter crews use (or mix) a number of frames of reference when exchanging information among themselves or transmitting to another vehicle: (1) ego-reference/spatial (e.g., a landmark is in front, to the right, or to the left of the pilot; the pilot should turn right or left); (2) ego-reference/clock position (e.g., a feature is at the observer's or recipient's 2 o'clock position); or (3) world-reference/compass heading (e.g., the pilot should look for a stream running North/South; the pilot should turn 20 degrees to a new heading of 280 degrees).

Ego-referenced directions are the easiest to process; they require minimal mental transformation or interpretation. Clock positions are less intuitively obvious than right/left directions, although they provide more precise information. However, clock position may be ambiguous if the sender's and receiver's points of reference (i.e., head position) are significantly different. Furthermore, extracting spatial information given in a verbal form may require additional mental transformations. When giving ego-referenced directions, the originator of the message must mentally project himself into the point of view of the intended recipient, an activity that imposes additional cognitive demands and is subject to error. Spatial information that is world-referenced (i.e., to a numeric or verbal compass position) is more precise than other forms, and does not require that the sender or recipient project

14

themselves into another's ego-reference. However, steering commands referenced to compass position pre-suppose that the recipient knows the current heading. In helicopters, pilots may have no idea what their current heading is (they focus on the external scene, rather than the instruments).Thus,the navigator might couple an ego-referenced command (e.g., turn right) that requires minimal mental transformation with a world-referenced modifier (e.g., Turn right...Now you're heading due West) to improve the pilot's orientation.

In addition to the problems associated with the use of different reference systems, helicopter crews often operate in unfamiliar environments where crewmembers do not share a common knowledge base about the names and appearance of significant landmarks. Thus, information about these landmarks must be transferred on the basis of their physical appearance (e.g., a small round pond; a dry river bed; a saddle-back hill), rather than by name (i.e., Jones' farm; White Mountain; Route 50). Given the potential differences in personal experience, descriptive terms may also have very different meaning for different crewmembers. For example, what looks like a pond to one, may look like a lake to another. A 500 foot hill might look like a mountain to a mid-Westerner, while a pilot from Colorado might describe it as a small hill, and so on. Furthermore, lack of familiarity with local vegetation may make the description process particularly difficult; it is easier to identify a grove of trees by name than by their physical appearance.

Thus, the task of navigation for helicopter crews is quite different than it is for automobile drivers (whose current route is always visible and identified by road signs) or transport pilots (whose current route is displayed on an instrument and identified by an explicit value).

Vehicle Control

When flying at very low altitudes, helicopter pilots' vehicle control inputs are based primarily on visual cues extracted from the external scene. In this respect, their task is similar to that of automobile drivers (except that they must also regulate altitude). Since they do not have a visible route to follow, helicopter pilots regulate speed, heading, and altitude so as to maintain a safe speed (given their proximity to the ground and obstacles) and adequate clearance, while continuing to head in the general direction of their goal. Maintaining a specific altitude, speed, or heading is less important than remaining clear of obstacles. In addition, helicopter pilots must control not only the direction in which their vehicle is moving, but also its orientation (the tail rotor must not slew around and hit an obstruction to the side, rear or below the cockpit) and assure adequate clearance for the rotor blades (which extend beyond the width of the vehicle).

Because helicopter pilots must continuously move their heads and eyes to scan the environment to avoid obstacles and search for landmarks, the dynamic optical cues used for flight-path control are often viewed off-axis with respect to the direction of travel. This adds to the difficulty pilots encounter in using dynamic optical variables to regulate speed, heading, and course. Figure 6 presents the dynamic optical flow cues that might be available when a pilot is looking forward, 45 deg to the left or 90 degrees to the left. As you can see, the information provided is distinctly different.

Helicopter pilots estimate speed by interpreting dynamic visual cues in the environment or listening to the sound of the rotors. To estimate velocity from dynamic optical flow, however, they must also estimate their altitude; apparent speed depends on the pilot's height above the surface over

15

which he is traveling. Alternatively, pilots may check their airspeed by looking at the instrument panel or by verbal information given to them by the navigator. Helicopter pilots estimate and maintain vertical and lateral trajectories and clearance from obstacles by monitoring the environment. They do not rely on instruments to control lateral or vertical position or rate when flying NOE. Again, information in the visual scene (e.g., dynamic optical flow, edge rate, and perspective transformations of features in the environment) is useful for detecting the effects of disturbances (e.g., winds) and pilot-induced deviations. As is the case with fixed-wing aircraft, control inputs affect more than one parameter. Thus, helicopter pilots must integrate their control activities to achieve a desired change. Because visible changes in optical variables may reflect changes in more than one axis, helicopter pilots must interpret the meaning of such changes, rather than responding to them directly (as they might when relying on instruments).

When flying with night vision devices, minification or magnification created by improper calibration or positioning of the lenses may impair the accuracy with which pilots can obtain dynamic motion cues. Furthermore, the reduced field of view that they provide (in current systems, the field of view is only 40 deg), limits the availability of peripheral motion cues. When using a helmet display of infrared imagery (such as provided in the AH-64 Apache helicopter), pilots face yet another problem. The sensor is located 3 ft below and 10 ft in front of the pilot's eye position. Thus, the pilot's visual reference is displaced. This produces systematic distortions: The vehicle appear to be moving faster and lower (because the sensor is closer to the ground than the pilot's usual visual reference) and obstacles seem closer than they are (because the sensor is forward of the pilot's usual visual reference). Since the display is presented on a monocle positioned in front of the pilot's right eye, binocular rivalry may be created by features visible in the external scene to the pilots' unaided left eye. Finally, symbology superimposed on the dynamic scene may interfere with the pilot's ability to detect subtle changes in the environment and create apparent-motion illusions.
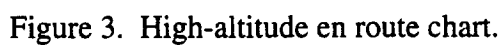
SUMMARY

Automobile drivers, airline pilots, and helicopter pilots use their eyes to obtain information for both vehicle control and navigation. The process of searching the external scene to find landmarks (for navigation) is intermittent and deliberate, while monitoring and responding to subtle changes in the visual scene (for vehicle control) is relatively continuous and "automatic." However, since operators may perform both tasks simultaneously, the dynamic optical cues used for vehicle control may be determined by the operator's direction of gaze for wayfinding. In some cases, the visual information acquired for one type of control activity may simultaneously provide useful input for another; when a helicopter pilot looks at the forward scene to avoid obstacles, information about rate of movement is also available from the flow of terrain past the vehicle. Conversely, the visual requirements of one control task may interfere with the requirements of another; when an automobile driver turns his head to look at a sign, his vehicle may drift out of its lane. Thus, in order to understand the use of dynamic visual cues for regulating vehicle motion, the simultaneous tasks of navigation and obstacle avoidance must be considered; operators do not just use their eyes to look for dynamic optical cues. Rather, they often look for landmarks or at potential threats, and coincidentally extract motion cues useful for vehicle regulation. Since the operator is no longer looking in the direction that

the vehicle is traveling, the optical relationships among cues in the visual scene may be somewhat misleading.

This chapter related the visual processes involved in vehicle control and wayfinding, contrasting the frames of reference and information used by automobile drivers, airline pilots, and helicopter pilots. The goal was to describe the contents within which different vehicle control tasks are performed and to suggest ways in which the use of visual cues for geographical orientation might influence visually guided control activities.

Figure 1. Helicopter low-altitude en route chart with iconic symbology.

Figure 2. 3-D conceptual chart of Los Angeles.

Figure 3. High-altitude en route chart.

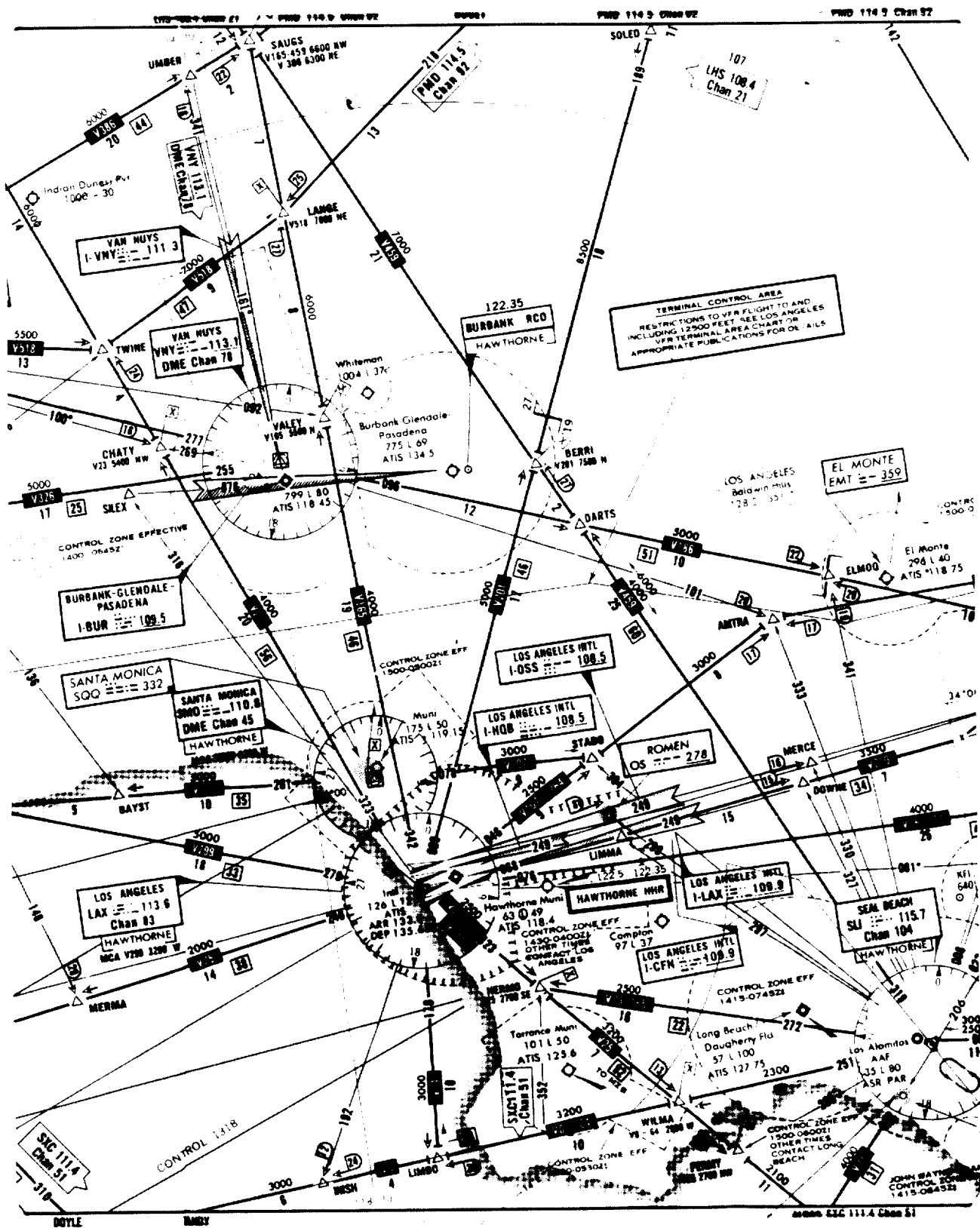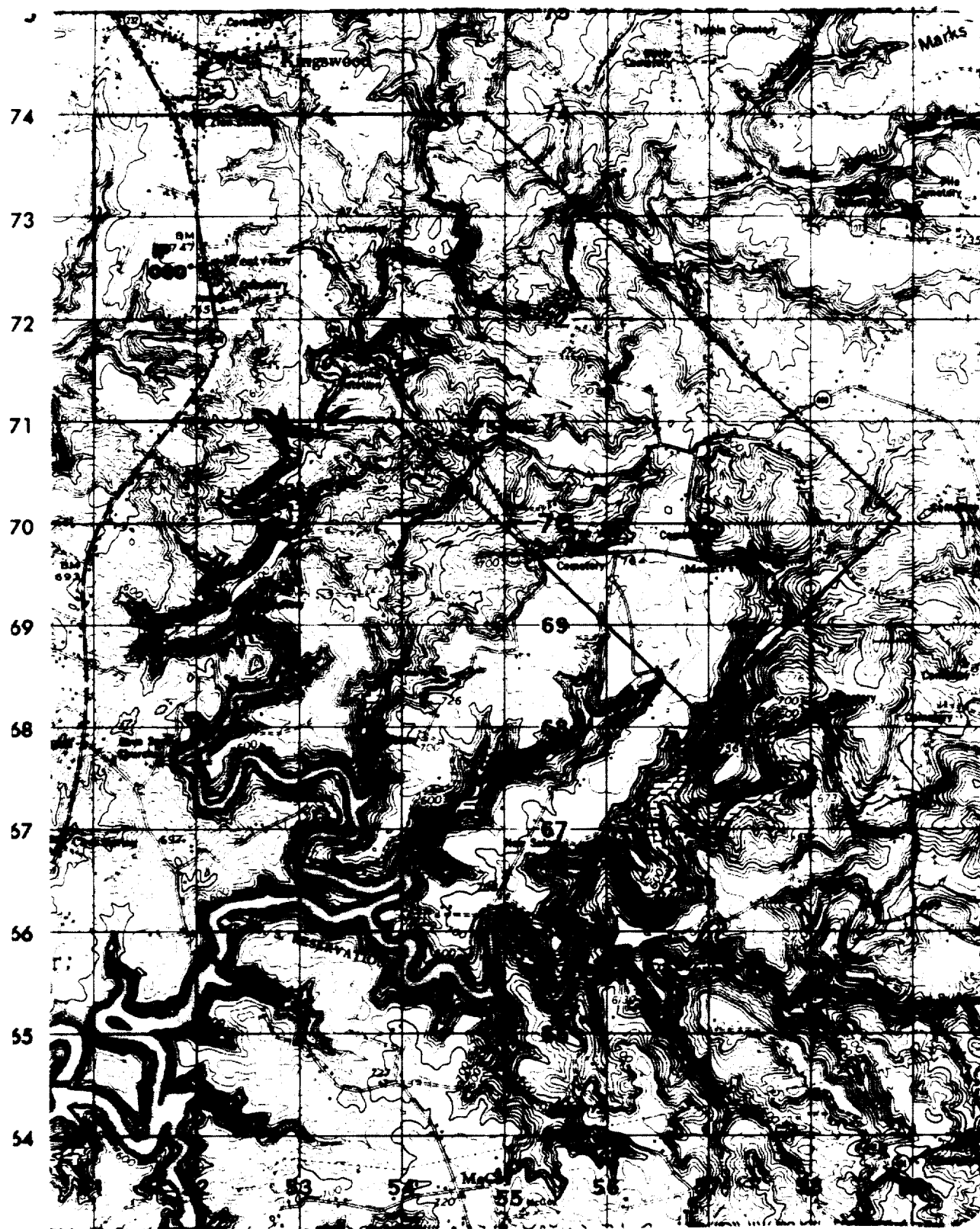Figure 4. Low-altitude en route chart.

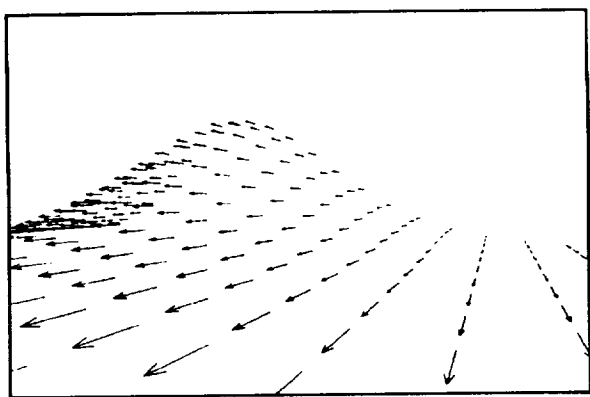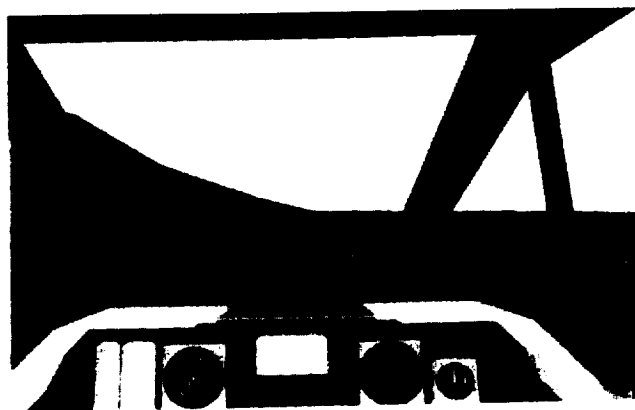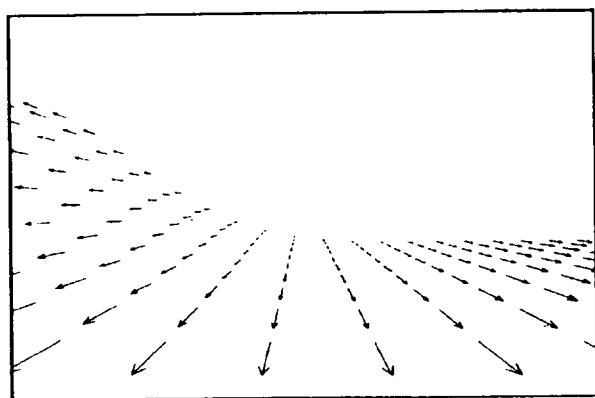Figure 5. DMA 1:50,000 tactical navigation chart.

Figure 6. Helicopter forward and left window views and flow fields.

THE DISPLAY OF SPATIAL INFORMATION AND VISUALLY GUIDED BEHAVIOR

C. T. Bennett
NASA Ames Research Center
Moffett Field, California

INTRODUCTION

The basic informational elements of spatial orientation are attitude and position within a coordinate system. The problem that faces aeronautical designers is that a pilot must deal with several coordinate systems, sometimes simultaneously. The display must depict unambiguously not only position and attitude, but also designate the relevant coordinate system. If this is not done accurately what will occur is, at the minimum, spatial disorientation, at the worst, catastrophe. This paper explains the different coordinate systems used in aeronautical tasks and the problems that occur in the display of spatial information used by pilots for aircraft control.

Pilot tasks and information sources. In order to successfully complete a flight mission, pilots traditionally have been taught to:

First--
　　Aviate,
Then--
　　Navigate,
Then--
　　Communicate.

Essentially, the first two of these are visually controlled tasks. The primary type of visual information used to accomplish these tasks will vary widely, depending on the task and the source of the information.

At one extreme, vehicle control tasks may be heavily dependent on visual information that is primarily sensory in nature. This might be the case if the primary goal of the control input is to regulate a specific aircraft state. Motion states are defined by a vehicle's three rotational (pitch, roll, and yaw) and three translational (longitudinal, lateral, and vertical) vectors. However, if the primary goal of a control input is ground track control (e.g., navigation), the pilot may rely primarily on cognitive synthesis of available visual information. The resulting knowledge may be in the form of perceptual or cognitive constructs. There are several other classes of visual information that are important to flight path guidance, but are related only secondarily to primary aircraft control and navigation. These include the display of radar weather returns, threat target locations, and traffic collision avoidance information.

Most of the non-visual sensory systems studied have been shown to have a significant impact on primary visual percepts. In the past, aeronautical display concepts typically have not taken advantage of the synergistic effects of polysensory inputs. This is the case despite the fact that polysensory displays may improve visual detection and recognition of optical events. Additionally, it may be that a given non-visual sensory system might more efficiently represent certain information. For example, it is possible that an auditory display may provide certain advantages in the display of spatial information. Accordingly, such displays merit systematic evaluation in an aeronautical setting.

SPATIAL INFORMATION DISPLAYS

As suggested by the aphorism presented earlier, the primary flight tasks are twofold: 1) control of translation and rotation of a craft, and 2) regulation of a craft's course. During flight through calm skies, these tasks apparently can be accomplished simultaneously. But, even the most novice of pilots soon learns that, during a flight, the mildest of winds can insidiously de-couple the control actions necessary to maintain orientation from those necessary to control the craft's course.

Decomposing these two tasks permits a relatively straightforward explanation of why, for the most part, a pilot has difficulty in performing them in parallel. The first task requires a reasonably solid understanding of aerodynamics, the science of the forces acting on a body in motion relative to air. The foundations of the second task lie in navigation, the science of determining position, course, and distance traveled. While pilots enjoy their amateur status as aerodynamicist and navigator, many find it nearly impossible to be both at the same time.

Spatial orientation and orienting usually refer to rather global tasks like determining attitude, position, and course. Early on in the history of flight, it was discovered that pilots are very poor at determining spatial orientation without the aid of reliable instrumentation. In fact, all of the primary flight displays were designed with only one purpose in mind: to maintain spatial orientation.

To be sure, certain flight tasks can be accomplished as accurately (or even better) by using the real world, perspective transformations as when the primary flight instruments are used. This statement is dependent, however, on the vehicle states. For example, at 100 feet there is a substantial amount of visual information generated by optical flow patterns. Such patterns can be used easily as visual cues. As a result, flight control based on the perspective transformations is possible. However, at 10,000 feet during straight and level flight there is little optical flow available; and, the pilot must rely on instruments for many flight control tasks.

The key to understanding the effectiveness of these displays is to realize how each of them supports the pilot in fulfilling the role of aerodynamicist or navigator. The key to the design of these displays is to understand how the information presented in each of them supports the pilot's ability to maintain spatial orientation.

THE PILOT AS AERODYNAMICIST AND NAVIGATOR

Aerodynamicist. The fundamental issue that produces de-coupling of aerodynamic and navigational control is the term "relative motion." As was pointed out earlier, aerodynamics deals with the motion of a body not just through air (although that is implicit), but motion of a body relative to air motions (wind). Knowledge of ground speed (which is calculated relative to the surface and is a navigation term) is unnecessary for the aerodynamic control of an aircraft. What a pilot must regulate is the rate at which air molecules pass over the wings. This information is displayed to the pilot by means of a sensor and gauge called the air speed indicator. However, air speed only represents part of the information set that is necessary to determine the rotational and translational motions of an airframe relative to the air mass.

All of the primary flight displays are specifically designed to provide some information concerning the six basic motion states of the craft. What complicates the design problem is that, in some cases, these displays provide information about orientation with respect to different coordinate systems. For example, the inertially referenced attitude indicator provides accurate information concerning pitch and roll relative to an earth fixed coordinate system. On the other hand, the airspeed (ram air) and altitude (barometric) indicators are referenced to the air mass coordinate system.

Suffice it to say at this point that understanding the translation and rotational motions of the craft "relative" to the motions of the air is necessary to avoid catastrophe. It follows that control inputs should first meet aerodynamic requirements. In fact, many accidents have resulted from a pilot's control inputs that are intended for navigational control without regard to the aerodynamic consequences. An example of this is when a pilot commands a bank to initiate a course change without considering the loss in lift that inevitably results when the craft rolls. A second, more dramatic example occurs when a pilot is low on a final approach course. A pitch-up command might appear to the novice pilot the simplest way to intercept the glide slope, and avoid landing short. But, a low approach in conjunction with a pitch-up command can be a deadly combination, for it results in increased drag, loss of lift, and loss of altitude. The FAA accidents classify such events as "controlled collisions with the ground."

Navigator. Generally, navigation is based on a pilot's ability to understand and control craft motions relative to true or magnetic north. The primary flight displays that are designed for navigation provide specific, although not necessarily complete, angular information about position relative to magnetic north or relative to some ground location. The pilot must take this angular information, convert it to longitude and latitude coordinates, and then plot the position on a chart. The positions, plotted over time, will provide the information necessary for accurate navigation (location, course, and distance traveled).

In addition, navigation is often considered to be a two dimensional task. (After all, charts are two-dimensional.) But, in fact, aeronautical navigation is three-dimensional (longitude, latitude, and altitude). The charting of a vertical flight path is necessary in order to establish cruise or descent profiles used for for calculating ground speed. And, just as lateral flight profiles are important for obstacle avoidance, vertical flight profiles are important for air traffic separation.

27

But, as was stated earlier, a pilot must first understand and deal with the motions of his craft relative to the air. In windy conditions, this pilotage canon will necessarily complicate the task of navigation. The classic example of such a case is when crabbing is required to make a direct translation between two ground locations. In the no-wind condition the craft heading (direction of the nose) is co-planar with the craft course (ground track). To maintain a rectilinear course on a windy day, the pilot must yaw the longitudinal axis of the craft (change heading) out of the plane of translation. When a pilot does this, the craft is no longer pointing (heading) where it is going (course direction).

The five coordinate system problem. It is important to recognize that there are two fundamental tasks that pilots must perform. One concerns getting from one point to another in the world. The other deals with keeping the craft in the air. However, it is also the case that for proper flight control, there are, in fact, five coordinate systems with which a pilot must deal simultaneously. They include three earth centered systems: inertial, magnetic, and polar. A fourth coordinate system is generated by the three planes normal and orthogonal to the relative wind. (The term relative wind is defined by the flow of air parallel to the craft's translational vector.) The fifth system is based on the longitudinal, lateral, and vertical axes of the craft. The challenge that faces a pilot is understanding the relationships among these coordinate systems. For proper flight control, they must be able to specify the impact of a simple control action on a craft's orientation in each of the coordinate systems.

A frequent and simple solution to the problem, but also a most dangerous one, is to ignore the way in which a single control input will be transformed through the different coordinate systems. The training a pilot receives emphasizes that control inputs directed toward a navigational goal will not necessarily assure aerodynamic stability. Often, however, a pilot does not learn that primary flight displays will not automatically sort out the interrelationships among the various coordinate systems.

CURRENT CONCEPTS IN SPATIAL INFORMATION DISPLAYS

Currently, there are two basic approaches to the graphical and numeric presentation of spatial information in a cockpit. One group, the primary flight displays, are the ones with which most people are familiar. The basic characteristic of such displays is that they present spatial information in an abstract format; for example, translational speed is displayed in the form of air speed or vertical velocity. The second general approach is called the contact analog display. It is designed to present a perspective, naturalistic representation of the crafts' motions that could then be easily related to abstract information in the primary flight displays. Typically, such a display will represent the craft moving over a ground plane.

The information that these displays present is very explicit concerning various vehicle states. However, the information in a given display is not necessarily specific to a coordinate system. The problems generated by this lack of specificity is discussed in the following sections.

Primary Flight Displays

Typically, primary flight displays present information about single aircraft states. Examples are air speed and vertical velocity indicators (translational rate displays), and the magnetic compass and directional gyroscope (rotational position displays). It should be noted, however, that some indicators, such as the turn coordinator, combine information about two states (yaw information and roll rate information).

Navigation-related displays also normally present a single dimension of guidance information [e.g., the VHF Omni Range (VOR) indicator presents bearing to a specific ground location relative to magnetic north, and the Automatic Direction Finder (ADF) presents this bearing relative to the nose of the craft]. However, like the turn coordinator, the localizer/glide slope display used for an instrument landing (ILS) combines information about two states (vertical and horizontal angular position). Although these displays primarily provide navigation information, they also provide indirectly attitude information that is used actively by the pilot. For example, if the pilot is monitoring heading by means of the directional gyroscope, any movement in the indicator specifies changes in roll and/or yaw.

Finally, since a pilot observes a display over time, the temporal dimension is present implicitly in all displays. While the time dimension is implicit, pilots explicitly use it to determine velocity or acceleration information (what pilots refer to as "trend" information).

Contact Analog Displays

The origin of the term contact display has its roots in the term contact flight. The latter has been given a specific usage by the FAA. It makes reference to a pilot's ability to fly and navigate by visual reference to the surface.

In the strictest sense, a contact flight display incorporates the perspective projection of a three dimensional model onto a picture plane. Typically, these displays represent a ground plane and a command path for a pilot to follow. In practice, the definition of a contact display is quite loose; examples of such displays have ranged from video displays to head-up-displays (HUD's).

The intent of contact flight displays was to take advantage of the eye's natural ability to sense and perceive motion in a perspective projection. Early in the history of these displays, questions arose concerning the design criteria for the field-of-view (FOV), field-of-regard (FOR), and resolution requirements. Little, if any, attention was directed to specifying surface texture element criteria.

Several studies have suggested that, for "normal" flight conditions, there are few differences in pilot control responses due to using contact or primary flight displays. Other studies suggest, that when the pilot is "stressed," performance with the contact analog display is better. However, caution should be exercised in generalizing from such studies due to the inadequate operational definition of the "stressor" variables.

What apparently draws engineers and designers to contact displays is the intuitive notion that if a naturalistic representation of the outside world can be presented to the pilot, performance will be

29

enhanced. This point of view is based on the assumption that a pilot can extract the information in a multi-dimensional perspective representation more efficiently than from a traditional single dimensional state variable display. It is assumed that a pilot can fly more accurately using dynamic perspective cues than using abstract informational displays. However, the conditions under which such assumptions may be true have yet to be defined.

Primary and Contact Flight Display Tradeoffs

The potential amount of information content in each of these two classes of displays is vastly different. As a result, the cost to the pilot in using them may also be very different. As indicated, a typical primary flight display presents a single dimensional state of the vehicle (e.g., airspeed). This display format has the benefit of being simple to read and interpret; but, several displays have to be read and integrated to acquire information concerning the overall state of the vehicle. This may not be a problem if the time required to use several single-state displays is minimal. Though it has not been well documented, experienced pilots can reportedly "read" an instrument panel at a glance, in a fashion analogous to someone who is learning to play chess. It has been argued that as one progresses from chess novice to chess expert, the essential skill that is acquired is the ability to perceive general patterns and the possible trends that might emerge. Apparently (and emphasis should be placed on the word apparently) pilots can perceive and determine multiple vehicle states with a single glance. It should be emphasized that this ability has not been demonstrated. It may well be that experienced pilots, particularly instrument-rated pilots, merely have a more disciplined and efficient instrument scan.

Conversely, it may not be most efficient to present multiple vehicle states simultaneously, as is done in contact analog displays. The notion here is that because a contact display is a representation of the real world, pilots would be able to use the information in the display as efficiently as they use the information in the real world.

There are three assumptions implicit in the supposition that the contact analog display format is more effective. The first is that we can in fact use efficiently the information in the real world. But, unfortunately, there are many ambiguities in the world scene that make motion sensing difficult. It may be that single dimensional primary flight displays are less ambiguous, and, thus, can be used more accurately. The second assumption is that pilots are sensitive to the graphical elements that are used to depict the real world. However, little research is available that specifies the visual cues in a graphical scene used by a pilot to control his virtual motion, and, more importantly, whether they are the same visual cues he would use in directly viewing the real world. A third assumption is that pilots rely on perspective cues to control translation and rotation. Under certain flight conditions, a pilot may simply rely on the two-dimensional information in a scene (e.g., image size).

Integrated Primary/Contact Flight Display

An alternative display concept is to integrate features of primary flight and contact analog displays into a single instrument. However, the benefits gained from such integration may be lost due to the added complexity.

30

Attempts to incorporate positional/rate information (in analog format) into contact flight displays have been limited. Examples of such attempts include "tunnel-in-the-sky" displays developed a decade ago. In such a display, a pictorial representation of changes in vehicle states are represented by simultaneously displaying slightly altered images of the same object, as if you were observing a cube from successively different viewing angles. This is akin to presenting sequential cartoon images closely in time, that are then assembled by the observer into a coherent motion percept.

Just as there have been attempts to integrate specific position/rate information into a contact display, there have been attempts to display plan-view navigation information into contact displays. Boeing is currently testing several of these displays. Other approaches have attempted to employ a "God's-eye-view" display of the craft's position.

However, trying to integrate features of these two display types will necessarily result in embedding even more information into the display. Whether this will facilitate information extraction by the pilot is another matter.

Cartesian versus Polar Coordinate Display Strategies

In developing display concepts, the designer has the freedom to specify the coordinate system (e.g., inertial) in which the information is presented. Freedom is also permitted in selecting the mathematical coordinate transformations used to specify position. The nature of the coordinate transformation depicted may influence the control strategy used by the pilot. Additionally, the designer is permitted freedom to "condition" the displayed information by a wide variety of filtering and lead/lag algorithms. Such techniques, while critical to design criteria for aeronautical information displays, will not be dealt with in this paper.

Cartesian coordinates. The assumption implicit in the design of most of the displays discussed is that the pilot has an internal representation of his motion through a Cartesian coordinate system. This assumes that pilots represent their space as if it has three intersecting planes which are orthogonal to each other.

This space is specified by three axes (x, y, and z) which provide distance metrics. A change in position is represented by a change in x, y, and z locations. To specify changes in rates of a vehicle, it is necessary to specify change in position over time in each of the three axes independently. While this may seem a bit obvious, the implication is that a single term cannot be used to describe something even as simple as approach speed on a glide slope because forward velocity must be computed independently of vertical velocity.

One potential problem a Cartesian based coordinate system display may generate is that it would direct the pilot to the "one-up-two-over" control strategy. That is, the display may lead the pilot to consider it is most efficient to control motions in different planes independently. For example, the standard ILS display in current cockpits shows angular deviation from the approach course and glide slope (horizontal and vertical planes). This causes many pilots to sequentially control either position on the approach course or glide slope. Such a response strategy may be contrasted to one in which a single control action is used to correct deviations in both approach course and glide slope.

31

Polar coordinates. Space also may be defined in terms of a polar coordinate system. Location in this space depends upon a single vector term [which is defined by its slant angle (a combined azimuth and elevation term) and the distance between the origin of the sphere and the reference point]. Changes in flight path angle and heading only require changes in a single value (slant angle). Changes in rate only require changes in the magnitude of the vector. Such a display may generate a ballistic control strategy based on craft dynamics.

Currently a vector display has been fielded for the control of hover location. In this case, there is a two dimensional vector (only x and y information is represented) that presents which direction and how fast the helicopter is moving away from a designated location depending upon the magnitude of its components.

An application of a Cartesian coordinate strategy to the same problem would present a surface with a dot moving around a specific reference location. Rate information would not be directly displayed, as it is in a polar coordinate display, but would have to be derived over time by the pilot.

Control strategies. The mathematical strategy that a designer uses to represent space may influence the nature of the control strategies a pilot uses. It may well be that different displays and/or control strategies may result in optimum performance depending upon the task. For example, when ballistic motions relative to the current location are sufficient (e.g., the hover-hold task), then a polar coordinate display may be optimal. On the other hand, when a pilot is flying close to the surface and needs to consider obstacle avoidance, a Cartesian-based coordinate display may be optimal.

SPATIAL INFORMATION DISPLAY CONCEPTS AND VISUAL ATTENTION

In any given display, there are often several sources of information. One goal of the display designer is to make it easier for the user to extract the information that is most highly correlated with an optimal response. Perhaps one of the most frustrating outcomes of display design is that the observer attends to information in the display that results in sub-optimal performance. This may occur because the "secondary" source of information is more compelling, or because the observer is more sensitive to it. For example, a perspective scene is generated to simulate translation over the real world. However, a pilot may not attend to the three-dimensional perspective transformations (which provide the mathematically optimal solution), but to two-dimensional motions of the surface texture elements against the edge of the screen.

One display design strategy is to physically co-locate information on the display surface (or even overlay information) so that the pilot can "simultaneously" assimilate both information domains. A classic example of this approach is the HUD. The intent, in part, was to overlay symbolic information on the real world scene, thereby reducing the amount of time it would take a pilot to integrate both information sources.

Several suppositions were made when this design strategy was conceived. One was that, because two information sources are proximally located, assimilation time would be reduced. This would be the case if the critical path component was movement of the eyes from one spot to another, and not

the time it took to sense and perceive the data source. A second assumption was one of "simultaneous assimilation." However, data existed even at the time when HUD's were first designed that suggested that parallel processing of diverse visual information may occur only under fairly limited circumstances.

The psychological literature is replete with studies on position perception. But, even now, little is known about how well an observer can perceive and control the motion of an object in three-dimensional space. Does the observer control motions in the x, y, and z planes in a serial fashion? Or, does the observer make a ballistic move between two locations, and only then check the error in the x, y, and z planes? Another way to state the question is the following: Does the observer treat each plane as a separate information dimension? An even more difficult question is how to tell which strategy a pilot is using.

This question illustrates one of the major issues in the design of visual displays; and it concerns the capacity of the visual system to parallel process (or, at least, multiplex) separate channels of multi-dimensional display. The problem of visual attention and display of spatial information often reduces to one of two issues. One concerns "tunneling" of visual attention. That is, information presented just a few degrees eccentric to the line-of-sight (LOS) may or may not be visually perceived. The other concerns the overlaying of visual information (the HUD strategy). What is not clear is whether or not a pilot can parallel process visual information that is overlaid, but is in two different planes. Stated in a more operational form, the question is: Can a pilot actually see and use information on the HUD while simultaneously attending to ground features?

NAVIGATION AND SPATIAL INFORMATION DISPLAYS

Typically, navigation displays are developed in isolation from the design of aerodynamic control displays because it is assumed that control and navigation displays are unrelated. Nothing could be farther from the truth. Unfortunately, designers have made this mistake; and, tragically, some pilots have made this same error and have killed themselves and their passengers. The following section describes the problem and discusses ongoing research efforts that address it.

Orientation and Multi-Coordinate System Registration

Relative wind and earth-relative coordinates. It was pointed out earlier that pilots are taught to attend to their aerodynamic tasks before attempting to solve their navigational problems. It was also described that, at times, a control movement made to solve one flight task is incompatible with solving the other. How a display might represent the impact of a control input on vehicle states in the different coordinate systems has received little attention.

Earlier, it was pointed out that the conventional wisdom of novice pilots often results in a pitch-up control input when his aircraft is low on a final approach course. The lack of understanding concerning the motion requirements for safely "navigating" in an air mass versus the motion requirements for navigating relative to a fixed earth position has unnerved many flight instructors.

Relative wind and magnetic north. A similar flight control problem was described concerning polar or magnetic transformations. The directional gyroscope has a compass rose that rotates as the aircraft yaws, while the number at the top of the display represents the magnetic heading. Due to the design of the display (it looks like a compass rose on a chart), and its dynamic characteristics (its motions are similar, but opposite, to the magnetic compass), a natural response is to treat the directional gyroscope as a navigation display. Unfortunately motions in this instrument also indirectly indicate vehicle yaw and/or roll in the air mass. In fact, during partial panel emergencies, pilots are taught to use the directional gyroscope as a substitute source for information that is normally displayed by the attitude indicator.

Again, a pilot must sort out the different coordinate systems. However, there is only one condition when the directional gyroscope provides accurate information about orientation in the relative wind coordinate system: when the two coordinate systems are in registration (aligned). Under many conditions the two coordinate systems are aligned closely enough that pilots can disregard the differences in the two coordinate systems. However, this is not the case in high performance aircraft. A separate instrument was designed to provide information about major changes in orientation of the craft in the relative wind coordinate system that may not be clearly represented in displays more closely related to the other coordinate systems. The instrument is called the angle-of-attack meter; and it has saved many lives.

Part of the point of presenting these two examples is to show that design problems associated with classical navigational questions should not be considered in isolation. Unfortunately there is very little information available concerning how a pilot might confuse motion information in a navigation display with motion information necessary for aerodynamic control.

Magnetic north and true north. There is another classic problem that falls into this multi-coordinate system problem; and, it has plagued display designers for years. It is the non-registration of the polar (north/south) and magnetic coordinate systems. Proper pre-flight planning will minimize problems a pilot might have in conceptualizing the relationships between these two coordinate systems. But, the fact remains that because of the difficulty in bringing these systems into registration conceptually during flight, that the unpracticed and unprepared pilot will avoid using the magnetic compass except in dire emergencies. (The problem is not only that the two norths are misaligned, but that the planes that form the magnetic axes are curvilinear.)

Orientation Within the Navigational Coordinate Systems

Plan versus perspective view. The display of position location is fundamental to navigation. To accomplish this accurately and unambiguously is the challenge of the designer. There are several issues that are important to the design of navigation displays, but will not be dealt with here in detail. These issues are primarily related to the iconic representation of the world and its features. But it is understood that such factors will undoubtedly influence the "cognitive display" of the world that the pilot generates.

The traditional approach to navigation display design has been to present a plan-view of the world as seen from above. In addition, there are some plan-view navigation displays that present a

side view of the scene. This is often done when accurate altitude control is critical. An example of such a navigational display is the standard instrument approach plate.

With the advent of high-speed, high-powered, small sized graphics displays (and terrain data bases), came the possibility of presenting three-dimensional representations of the terrain. But, as long as the position information is presented graphically, virtually the same questions remain for three-dimensional navigation displays as there were for the two-dimensional plan-view representations. What viewing angle should be shown? How should the surface be depicted? What is the best way to represent cultural and vegetative features?

Coordinate transformation. Determining exact position location on a chart from the world scene and the information in the navigation displays is another critical navigation task. The pilot must take the real world scene, match it with some graphical representation in the display (or chart), and determine its associated coordinates. The task of determining the graphical/navigational metric values is a constant source of problems for the pilot. It is created by the fact that all of the typical charts available to the pilot provide location information in degrees of longitude and latitude. However, in the cockpit, the information about position location is typically presented in relative angular units (degrees).

This transformation problem is not simply solved by cockpit displays which provide longitude and latitude coordinates of the craft's current location. The pilot still must look out the cockpit wind screen, identify an object, determine its relative angular bearing to his craft (remember there are no longitude and latitude lines in the real world), then compute the object's location (in degrees longitude and latitude) based on his present coordinate location.

Design Criteria and the Display of Spatial Information for Navigation

Examples have been presented which show how navigation display design can influence the use of spatial information, particularly as a pilot controls his orientation in the other coordinate systems. This problem must not be disregarded if an accurate, as well as safe, display is to be designed.

To accomplish the above, the designer must realize that display principles that seem quite appropriate on the ground, where a controller has to deal with only two axes, may not be appropriate in the air, where there is not only another axis to deal with, but also additional coordinate systems. The challenge here is to track the impact of control actions in all five of the coordinate systems, and to avoid displaying those orientation changes which will lead a pilot astray, while he is acting as an aerodynamicist or as a navigator.

VFR/IFR TRANSITIONS: A MODEL FOR DISPLAY EVALUATION

Operational definitions versus operational relevance. In the preceding discussion, some of the aerodynamic and navigational tasks that face a pilot were outlined. These problems were presented in terms of the coordinate systems with which a pilot must deal, and some of the characteristics of current spatial information displays. The development of a coherent design criteria for the display of

35

spatial information is an imposing challenge to the perceptual scientist. To accomplish this, both the nature of spatial information, as well as vehicle control strategies must be better described. Such questions as the following need to be asked: How are points in space localized? Are different "mathematical" strategies used by pilots to specify translation and rotation in space? For that matter, how are vehicle dynamics understood by the pilot? In what manner are vehicle dynamics used by the novice pilot versus the veteran captain?

To formalize these questions for empirical scrutiny necessitates abstraction of the basic perceptual principles utilized during flight. The danger that lies in this process is that the resulting "operational definitions" (from an experimental methodology perspective) will not be "operationally" relevant from an aeronautical perspective. In an effort to minimize the problems that might occur during the transition from laboratory to cockpit, the following operational problem is parsed to clarify some of the relationships among perceptual constructs and pilot tasks.

Experimental model. Transitions between Visual Flight Rules (VFR) and Instrument Flight Rules (IFR) flight are considered to be some of the more formidable flight tasks. However, the specific nature of the difficulty is unclear. In the case when a pilot must transition from Instrument Meteorological Conditions (IMC) to Visual Meteorological Conditions (VMC), several sensory, perceptual, and cognitive tasks must be accomplished. While flying in IMC, the information from several one dimensional craft state displays must be integrated into a cognitive representation of position and attitude in space. This cognitive representation must include world features that will be encountered as the weather transition is made. Problems in disorientation may develop if the cognitive representation of the world does not match with that actually encountered.

Additionally, spatial disorientation (misperception of orientation and position in at least one coordinate system) may occur as the result of loss of the perception of vection. As a pilot transitions into the clouds, optical flow cues, which normally produce a sensation of vection, may not be present. The two-dimensional primary flight displays do not generate a sensation of vection. As a result, control inputs, which were associated with vection while flying in VMC, are suddenly dissociated from typical visual motion cues.

As the craft motions take it into VMC during an approach, the pilot will look out the cockpit window and his control motions will be influenced by the perspective transformations that are taking place in the world. The gain of the information in the world may be different than the gain of the primary flight displays. The pilot must adjust to differences in scale, format, and information content. Due to the total perspective transformation taking place, vection may be experienced. The pilot must adjust to this as well. As the pilot begins to recognize cultural and topographic features in the world, comparisons to the cognitive map he made during IMC flight will be made. These differences must be accommodated as well, and usually in a very short time period.

The most important problem that faces the pilot at these transitions are the extraction of position and attitude from uni-dimensional displays and the rapid mapping to the multi-dimensional world "display." Understanding the VMC/IMC transition process may well serve as a model for understanding the differences in performance when using primary flight displays versus contact analog displays. In addition, it may aid in the development of design strategies for representing different

coordinate systems. And, perhaps most importantly, a model such as this may serve as the basis the development of a design criteria for spatial information displays in aircraft.

SUMMARY

Visually guided control of an aircraft is dependent upon a pilot's understanding of his location in any one of several coordinate systems. Such coordinate systems may be relative to the earth, the craft, or the pilot himself. To control an aircraft within these systems, the pilot must understand their spatial relationships to one another, as well as the control laws and craft dynamics which may be specific to a given coordinate system.

To develop spatial information displays for a pilot, a designer must consider the (1) aerodynamic and navigational coordinate systems within which a pilot must control his aircraft, (2) the control task required of the pilot, (3) the mental model that defines the control space, and (4) how the pilot transitions from one coordinate system to another.

PERCEIVING ENVIRONMENTAL STRUCTURE FROM OPTICAL MOTION

Joseph S. Lappin
Vanderbilt University
Nashville, Tennessee

This technical report is for a Workshop on Visually Guided Control of Movement sponsored by NASA Ames Research Center, Rotorcraft Human Factors Branch, June 26 - July 14, 1989.

INTRODUCTION

Guiding a plane or helicopter over a natural terrain cluttered with objects of varying size, shape, position, and altitude requires extraordinary spatio-temporal coordination of the pilot's motor actions with optical information about the structure of the environment. Even though such visual-motor coordination is a commonplace achievement of human perception and action, the effort to understand this capability constitutes one of the main frontiers of contemporary science.

If the quantity of optical information utilized by a pilot in nap-of-the-earth flight, for example, is measured in terms of the number of bits per second per pixel that must be displayed by a high-fidelity, wide-field, realistic simulation by a computer graphics imaging system, then this quantity of information may approach roughly 10 billion bits per second - far beyond the capacity of state-of-the-art technology for acquiring or controlling optical image data. Human pilots, moreover, are able not only to visually acquire such optical information but to transform it in real time to coordinate the six-dimensional trajectory of an aircraft with the rapidly changing constraints of the surrounding environmental scene.

In fact, however, visually acquired optical information cannot yet be quantified. Despite extensive efforts and impressive progress in many relevant areas of science and technology over the past 25 years or so, we still lack a clear understanding of precisely what optical relationships constitute visual information. We cannot yet be certain exactly what properties can in principle or do in fact enable the real-time visual perception of 3D environmental structure.

Generally speaking, one of the most important sources of optical information about environmental structure is known to be the deforming optical patterns produced by the movements of the observer (pilot) or environmental objects. The visual salience and effectiveness of the information provided by such optical image motion has been amply documented by a large body of psychophysical research, by research on computer vision and robotics, and by a considerable body of experience in controlling flight in both real and simulated aircraft. As an observer moves through a rigid environment, the projected optical patterns of environmental objects are systematically transformed according to their orientations and positions in 3D space relative to those of the observer. The

detailed characteristics of these deforming optical patterns carry information about the 3D structure of the objects and about their locations and orientations relative to those of the observer.

The purpose of this paper is to examine specific geometrical properties of moving images that may constitute visually detected information about the shapes and locations of environmental objects. The basic theoretical ideas are the following:

(1) Optical information about environmental structure consists of two qualitatively different types of geometrical relationships which provide information about two different characteristics of environmental structure;

(2) First, information about the intrinsic geometric shape of environmental objects is primarily information about the differential structure of surfaces.

(3) Optical information about the differential structure of environmental surfaces is provided by local properties of the differential structure of moving images of the surfaces. In principle, this local image structure is sufficient to specify the metric structure of a local surface patch (up to a scalar), independent of other information or assumptions about the egocentric distance or orientation of the object relative to the observer.

(4) This information about local surface structure is based mainly on a rotation of the object relative to the observer (around some axis that does not pass through the observer's viewing position). The angular magnitude of this transformation provides a unique one-parameter transformation by which vision can represent the image transformations produced by the motions of environmental objects relative to the observer.

(5) Second, in contrast, the egocentric distance of an object, the distances between separated objects, the orientation of a given surface, and the observer's own location and motion within the environment all involve a qualitatively different aspect of the geometrical structure of the environment, specified by a different geometrical characteristic of the images. These geometrical properties reflect locations and orientations within an abstract 3-D Euclidean framework defined independently of the objects and motions within the space.

(6) Optical information about the structure of this abstract 3-D framework is defined by global properties of the images, specified by the (six) parameters of the perspective embedding of the 2D retinal image into Euclidean 3-space.

THE CONVENTIONAL CONCEPTION OF SPACE

"Space" is commonly regarded as an extrinsic framework in which Euclidean distances may be defined independently of the objects contained in the space. The sizes and shapes of objects, the distances between objects, and the velocities of objects moving in the space are usually considered as defined in relation to this framework, independently of the objects themselves. Thus, the spatial structure of optical data patterns on the retina has usually been represented in relation to the

40

anatomical arrangement of the photoreceptors - in a 2D space. Similarly, the structure of the 3D environmental space that contains both the observer's retinae and other environmental objects has been regarded as given a priori, independently of the observer and of other objects and events.

This conception of space has several important consequences: First, according to this geometrical representation, the problems of perceiving the structure of objects, their locations and orientation within 3D space, and the observer's own position and motion within this 3D space are all variations on the same general problem - namely, the problem of reconstructing or inferring a 3D environmental space from the 2D retinal optical patterns. As we shall see, however, two different aspects of environmental structure are probably described by two different geometrical characteristics of the retinal optical patterns.

Second, the 2D Euclidean distances on the retina cannot be isomorphic with 3D Euclidean distances in the environment. The mapping from the 2D Euclidean distances on the retina into the 3D Euclidean distances in the environment is necessarily ill-defined. Hence, computational solutions to the problem of recovering the 3D environmental framework require additional extra-retinal information, prior assumptions about natural constraints on the structure of environmental objects and events, and/or processes involving logical inductions and heuristics. Additional spatio-temporal information associated with movements of the observer and objects offers potentially important constraints on computational solutions of this problem, but this additional information is not sufficient to remove the fundamental limitation inherent in the mismatch in dimensionality of the retina and the environment.

Third, descriptions of the retinal optical data patterns are constrained by representing their spatial organization only in relation to the anatomical arrangements among the retinal photoreceptors. Thus, optical patterns are often represented as scalar fields - values of luminance at given spatial positions. Distances and other geometrical relations in the optical patterns are implicitly assumed to be defined only by reference to the retinal anatomy. Another common representation of the optical patterns is as a vector field - a binary relational structure where each vector is specified by two parameters, corresponding to its length and orientation or to the positions of its end-points. The optical velocity field is an example of such a binary relational structure, where the vectors correspond to successive space-time positions or directions and velocities of individual moving points. Because this geometric relational structure of the optical data patterns themselves is quite primitive, the visual computational processes required to recover the geometrical structure of environmental objects have necessarily been complex, time-consuming, and unreliable. For example, a difficult first computational step has been thought to involve solving the so-called "correspondence problem" - matching the spatial position of each point in one image with its corresponding position in a following image. The nature and difficulty of this problem, however, stems from representing the retinal spatial patterns as sets of points, without regard for the intrinsic geometric order of the optical patterns.

The justification for these geometrical representations of the optical patterns, however, has been based on presumption and convention rather than empirical evidence or theoretical analysis. We now examine an alternative representation of the geometric information in optical patterns which significantly simplifies the computational requirements for processing this information.

41

INTRINSIC GEOMETRY OF SURFACES AND THEIR IMAGES

The geometry of vision is simpler when described in terms of the intrinsic structure of surfaces. Surfaces are connected sets of points in 3D space, but they are just 2D manifolds - positions and changes in position on the surface can be described by just two independent parameters. (A "manifold" is a mathematical structure that is differentiable almost everywhere.) So long as one restricts attention only to points on the surface, geometrical relations in the abstract empty space outside the surface can be ignored as irrelevant to the geometry of the surface itself. Now the retinal images of surfaces are also 2D manifolds. Thus, the geometrical correspondence between surfaces and their retinal images is much closer than that between arbitrary collections of points in 3D Euclidean space and their perspective images on the 2D retinal surface.

The remarkable and important fact is that the differential structures of natural surfaces and their retinal images are isomorphic.[2] (In more technical jargon, we can say that the two structures are "diffeomorphic" - meaning that the mapping from the surface onto its image is one-to-one and differentiable and the same is true for the inverse mapping from the image onto the surface.) This isomorphism means that the differential structure of the retinal images of a surface provide rich and precise information about the differential structure of the environmental surface. This isomorphism holds for images defined by texture, motion parallax, and stereoscopic disparity; and even though it does not actually hold for images defined by illuminance, due in part to the conjoint influences of the directions of illumination and of gaze, the illuminance gradients do provide detailed information about the differential structure of the surface. It follows that the images defined by these various types of optical properties are also isomorphic with one another.

Particularly informative characteristics of the differential structure of a surface are given by its critical points; and a corresponding characterization in the image is provided by the critical values which are the images of the critical points. These critical points are isolated points and curves at which the differential map from the surface onto its image decreases in dimensionality from two to one dimension - at minima and maxima of the height of the surface (at peaks and valleys), at water sheds and water troughs, at parabolic lines or inflections where the curvature changes sign between regions of convexity and concavity, at saddle points, and at the discontinuities associated with sharp corners, where the derivatives of the surface vanish. (An additional set of discontinuities in the image corresponds to occluding and bounding contours where the surface is smoothly curved but the image is discontinuous. The surface positions corresponding to these image discontinuities do not remain constant as the object rotates relative to the observer.) The spatial pattern of these critical points provides a type of skeletal framework describing the topological structure of the surface independent of its orientation relative to the observer. Even in regions of the surface which appear and disappear from view due to changes in occlusion, the pattern of these catastrophic image changes is quite systematic and carries considerable qualitative information about the structure of the surface (see Koenderink & van Doorn, 1976).

IMAGES OF SURFACES DESCRIBED BY COORDINATE TRANSFORMATIONS

The mapping of the differential structure of a local surface patch onto that of its image can be very simply described as a linear coordinate transformation. This linear approximation of the perspective optical projection holds for "infinitely small" surface patches that may be locally approximated by a tangent plane at that location. Thus, the mapping of spatial relations on the surface onto those of its image can be locally described as a linear transformation of the two coordinates of the tangent plane at that location - from the intrinsic surface coordinates onto the retinal coordinates. Because the relative orientation of surface patches on the object and the image changes with the curvature of the surface and with the orientation of the object in the observer's visual field, the parameters of these coordinate transformations vary smoothly over the surface of the object.

Suppose that O2 represents the 2D manifold of the object surface, and that R2 represents the 2D manifold of the observer's retina. Then the linear map $v = O^2 \rightarrow R^2$ is locally specified by the following Jacobian matrix of partial derivatives:

$$V = \begin{bmatrix} \partial r^1 / \partial o^1 & \partial r^1 / \partial o^2 \\ \partial r^2 / o^1 & \partial r^2 / \partial o^2 \end{bmatrix}$$

Thus, suppose that $[dO] = [do^1, do^2]^t$ is a 2×1 column vector that describes an infinitesimal displacement on the surface in terms of two intrinsic coordinates on the object surface, and suppose that $[dR] = [dr^1, dr^2]^t$ is a corresponding description of the image of this vector in terms of the intrinsic coordinates of the retina. Then the transformation between these two coordinate systems produced by the optical projection from the object to its image on the retina is given by the linear equation

$$[dR] = V [dO] \tag{1}$$

and the inverse map is given by

$$[dO] = V^{-1} [dR] \tag{2}$$

where V is the Jacobian matrix given above. (The form of this equation is independent of the specific coordinate systems used to specify positions on the two manifolds. The coordinates need not intersect at right angles nor even be straight lines; they need only be differentiable and to provide a unique specification of each position on the manifold. The generality of this formulation seems especially relevant to vision, where no specific coordinate system can be assumed beforehand for any given environmental surface, and where the visually effective coordinates of the retina are not known.)

The coordinate transformation specified by the Jacobian matrix V may be understood as a local description of the retinal image. The parameters of this transformation need not be computed from more elementary data; these parameters constitute a representation of the image itself. The four

43

parameters simply quantify the local densities of the two retinal coordinates relative to each of the two object surface coordinates.

A principal characteristic of this representation is that it is a four-parameter description of the local relational structure of the image. Thus, these four parameters are necessary and sufficient for describing the local structure of the image.

Accordingly, the elementary optical image predicates for perceiving environmental structure from motion consist of the temporal deformation patterns of these four spatial parameters. Although the complexity of this relational structure exceeds that of the scalar or vector fields which are often used for describing such images, this greater complexity reduces the ambiguities associated with the so-called correspondence problem in matching component elements in successive images.

THE METRIC STRUCTURE OF A SURFACE FROM CONGRUENCE UNDER MOTION

The metric[3] structure of an arbitrary surface—providing quantitative measures of lengths, angles, and areas on the surface rather than in abstract empty space—involves an embedding of the surface into Euclidean 3-space. That is, perception of the intrinsic geometry of a surface involves three separate coordinate systems for describing any given infinitesimal displacement on the surface: intrinsic coordinates on the object's surface (O^2), retinal coordinates of the image of the surface (R^2), and Euclidean 3-space (E^3). Thus, we employ three differentiable mappings between three separate manifolds: $O^2 \rightarrow E^3$; $O^2 \rightarrow R^2$; $p: R^2 \rightarrow E^3$. A standard formula in differential geometry, the "first fundamental form," specifies the metric structure of a local surface patch based on its "natural" embedding, n, from O^2 into E^3. By using the chain rule for partial derivatives, we can express the natural embedding n as a composition of the functions v and p – i.e., $n = p \cdot v$ – and this leads easily to an expression for the metric structure of the retinal image of a local patch on the surface on an environmental object.

These relations are easily characterized by matrix equations. Let $[dX] = [dx^1, dx^2, dx^3]^t$ be a 3×1 column vector which specifies the lengths of a given displacement on each of the three orthogonal axes of E^3. Note first that the Pythagorean formula for distance can be written in matrix form as

$$ds^2 = [dX]^t[dX] = \sum_{k=1}^{3} \left(dx^k\right)^2 \qquad (3)$$

Next, we express this equation in terms of the intrinsic coordinates of the object surface. Let $N = \left[\partial x^k / \partial O^i\right]$, $k = 1,2,3$ and $i = 1,2$ be the 3×2 Jacobian matrix which transforms the description of a local surface patch from the O^2 into the E^3 coordinate system. Thus, $[dX] = N [dO]$. Now we substitute into the Pythagorean formula to express the quantity ds^2 in terms of the object surface coordinates:

44

$$ds^2 \quad = \quad [N\,[dO]]^t\,[N\,[dO]]$$

$$= \quad [dO]^t\,N^t\;N\,[dO] \qquad\qquad (4)$$

$$= \quad [dO]^t\,G\,[dO]$$

$G = N^tN$ is a symmetric 2×2 matrix whose entries are the metric tensor coefficients for the local surface patch,

$$g_{i\,j} = \sum_k \left(\partial x^k / \partial O^i\right)\!\left(\partial x^k / \partial O^j\right)$$

As may be seen there are three independent parameters in this matrix, g_{11}, $g_{12} = g_{21}$, and g_{22}. The values of these parameters remain invariant under rotational transformations of the E^3 coordinate system in which the object is described.

Now we employ the chain rule to find the metric tensor coefficients for the retinal image of the surface patch. By the chain rule we have $N = PV$, where P is the 3×2 Jacobian which embeds the retinal coordinates of the given surface patch into E^3, $P = \left[\partial x^k / \partial r^a\right]$, with $a = 1,2$. Thus, we have

$$G \quad = \quad N^tN = [PV]^t\;[PV]$$

$$= \quad V^t\;P^t\;P\;V$$

$$= \quad V^t\;P^*\;V$$

and

$$ds^2 \quad = \quad [dO]^t\;V^t\;P^*\;V\;[dO] \qquad\qquad (5)$$

where $P^* = P^tP$ is a symmetric 2×2 matrix with entries

$$p_{ab} = \sum_k \left(\partial x^k / \partial r^a\right)\left(\partial x^k / \partial r^b\right)$$

In this construction of the metric tensor coefficients, $G = V^tP^*V$, the components of V are directly specified in the retinal image, and the three parameters of P^* are unknown free parameters which constitute the metric tensor coefficients for the retinal image of the surface patch. The values of P^* are not determined by a single image but can be estimated from optical information associated with movements of the object or observer.

A principal hypothesis in the present theory is that the perceived metric structure of retinally imaged surfaces is derived from invariance of the shape of the surface under rotational motions. This hypothesis contrasts with the more common approach of deriving the local surface structure from the global structure of the space it inhabits—e.g., from the relative depth values of neighboring regions

of the surface. Here, the surface structure is regarded as fundamental, and the structure of the space containing the object is derived from the isometries induced by the motion of the object.

If V and P are respectively the "visual" and "perspective" coordinate transformations for an initial image of a given surface patch, then suppose that U and Q are the corresponding coordinate transformations for a second image of the same surface patch seen from a different observational position. Because the metric structure of this surface patch remains invariant under motions in E^3 (rigid rotations as well as bendings; translations of course have no effect on the differential structure), then we have

$$G = V^t \; P^* \; V \; = \; U^t \; Q^* \; U \tag{6}$$

The parameters of the visual transformations V and U are given directly by the two successive images of the given surface patch, and the parameters of the metric embeddings P^* and Q^* are unknown parameters which must be found as solutions of these equations. These two sets of perspective parameters represent six unknown parameters, for each of the two images. Unfortunately, the values of these six parameters are not determined by this matrix equation since it involves only four independent equations.

If the values of Q^* can be expressed as a one-parameter transformation of the values of P^*, say $Q^* = f(P^*)$, where $f( )$ is the desired one-parameter transformation, then we would require only four independent parameter values as solutions for the four independent equations. The one-parameter transformation that yields these solvable equations is a rotation which moves the surface patch over a surface of revolution (see Guggenheimer, 1963, pp. 272-273). Thus, for example, if the 3D surface is a sphere rotated around an axis through its center, the perspective and metric embeddings, P and P^*, of the image of any given surface patch on the sphere would be altered by such a rotation of the sphere, but the transformation of these perspective and metric parameters would be specified simply by the angle of this rotation.

Evidence that vision can indeed obtain sufficient information for perceiving the metric structure of a spherical surface from just two successive views which differ by such a rotational transformation was obtained by Lappin, Doner, and Kottas (1980) and Doner, Lappin, and Perfetto (1984). The displayed surfaces were seen as spherical even though the perspective projection used to display the surface seriously violated the normal perspective for 3D objects seen at that viewing distance, as if the object were seen from a position much closer than that at which it was actually presented. Related results were also reported by Lappin and Fuqua (1983), who found that observers exhibited "hyperacuity" for perceiving the center of the length of an imaginary line segment specified by three collinear dots rotated (about a nondisplayed collinear point) in a plane slanted by randomly varied amounts from the fronto-parallel plane. The surface of revolution in this case was a plane specified by the space-time trajectory of the rotating line segment. As before, this spatial discrimination performance was shown to be unaffected by the degree of polar perspective projection used to display the rotating line segment. Similar findings have also been obtained by Lappin and Love (in preparation) for discriminating the shapes of elliptical forms which were displayed stereoscopically on a plane slanted in depth by varying amounts. Large and variable magnifications of the stereoscopic disparities of the shapes were found to have little or no detrimental effect on discriminations of small differences in the relative shapes of these forms when they were rotated, although the shapes were

46

essentially indiscriminable when they were stationary. In general, then, the metric structure of these spherical and planar surfaces of revolution seems to have been accurately perceived, independently of the naturalness of the perspective with which they were displayed. Evidently, the perceived metric structure of these forms and spaces was enabled by the rotational transformations of the images of these forms.

Let us now examine the potential geometrical basis for this perceptual achievement. Suppose that PV and QU, respectively, are the perspective embeddings into E^3 of the first and second retinal images of a given surface patch, and suppose that these images are related by a rotation in E^3. Thus, we have

$$Q\ U = F\ P\ V \tag{7}$$

where F is a standard 3×3 rotation matrix. F is determined by three parameters, corresponding to the magnitudes of rotation around each of three previously given orthogonal axes. Any given momentary rotation occurs in only a single plane, however; if one of the orthogonal basis vectors happens to be perpendicular to this plane, then the rotation will have no effect on metric relations in that axis, and the metric relations in the two axes of the plane of rotation will trade off against each other.

Now if we wish to determine only the metric relations of the surface patch, ignoring the specific orientation of the surface relative to some other extrinsic reference system, then we can choose the basis vectors of E^3 so that one is perpendicular to the plane of rotation and the other two lie in the plane of rotation. Thus, the magnitude of rotation can be specified by a single parameter value, and the transformation can be described by a 2×2 matrix. We designate this restricted 2×2 matrix by F_2. Similarly, the changes in the perspective embedding parameters are also restricted to only two of the three axes of E^3, and the equations for the coordinate transformations produced by the rotation can also be described by 2×2 perspective matrices, say P_2 and Q_2. Thus, we can now rewrite Eq. (7) as the following equation involving only 2×2 matrices:

$$Q_2\ U\ =\ F_2\ P_2\ V \tag{8}$$

Because the matrices P_2 and Q_2 each now have an inverse, we can rearrange terms in Eq. (8) to represent the observed image deformation given by V and U in terms of an angular rotation in Euclidean coordinates between P_2 and Q_2:

$$U\ V^{-1}\ =\ Q_2^{-1}\ F_2\ P_2 \tag{9}$$

The left side of Eq. (9) specifies the observed image deformation defined by the two successive images V and U, and the right side is the representation of this deformation as a rotation in E^3. This matrix equation is composed of four independent quadratic equations in four independent parameters. Each of the terms on the left evaluates the relative magnitudes of the partial derivatives involving one of the two retinal coordinates for the second image of the surface patch relative to one of those for the first image of the same surface patch. The corresponding entries in the combined 2×2 matrix on the right side of Eq. (9) evaluate the relative magnitudes of partial derivatives that quantify the embedding of the same pair of retinal coordinates into the two Euclidean coordinates of the plane

47

of rotation. As the retinal coordinates for the image of a given surface patch expand (contract) in the second image relative to the first image, then the Euclidean embedding of the retinal coordinates contract (expand) in inverse proportion from the first to the second image, so that the metric embedding of the object coordinates of the given surface patch remain constant. (A more detailed presentation of the relevant equations is given by Lappin, in press.)

ROTATION AS THE BASIC TRANSFORMATION FOR PERCEIVING STRUCTURE FROM MOTION

One of the principal hypotheses implicit in these equations is that the angular magnitude of rotation in depth constitutes a fundamental relationship for visually perceiving the transformation between successive retinal images of a moving object. This representation is geometrically valid only in a restricted subset of cases, however - where the trajectory of the surface patch in Euclidean space-time is a surface of revolution. Most object surfaces and most trajectories do not really satisfy this condition. Even when an object rotates, the sequence of positions of most surfaces occupies a volume of space rather than a surface - e.g., consider a rotating cube or a sphere rotating around an axis that does not pass through its center. Accordingly, the perspective embedding of the images of the surface into E3 necessarily varies over time from one image to the next. Moreover, because this volume constitutes a three- dimensional rather than a two-dimensional manifold, the projective visual mapping of this manifold onto its retinal images is no longer diffeomorphic and no longer has a well- defined inverse.

Despite these apparent difficulties, the trajectory of an infinitesimal surface patch on a rotating object usually does approximate a section of a surface of revolution for at least a brief interval of time. The accuracy of the approximation improves as the area of the patch and the interval of time are reduced. For the "infinitesimally" small local patches on which the metric tensor is defined, the thickness of the volume is negligible in relation to the other two dimensions of the surface. Moreover, the neighboring patches on the object's surface have trajectories described by the same angular rotation, differing smoothly only in their radial distances from the axis of rotation. It is the differential structure of these radii of rotation that is the goal of these visual analyses, not the angular rotation parameters as such.

A second apparent limitation of this geometric approximation is that the group of motions in E^3 includes other motions besides rotations in depth. Translational movements of the observer as well as those of objects are common visual events and these transformations of the optic array are potentially important sources of visual information about 3D structure. Two different classes of such translational transformations are pertinent: (a) translations approximately parallel to the direction of gaze, which produce "looming" or divergence of the optical images, and (b) translations approximately perpendicular to the direction of gaze, yielding the classical "motion parallax" cue. Each of these cases is considered below. To anticipate, present evidence suggests that (a) the optical divergence patterns produced by approaching or receding objects are visually ineffective as information about surface structure (though potentially useful as a source of information about egocentric distance); and (b) the motion parallax patterns associated with translations roughly perpendicular to the direction of gaze are visually perceived as if produced by rotation rather than translation.

48

Translations that coincide with the direction of gaze produce optical flow fields characterized by divergence or "looming". The trajectory of any given point in the retinal image flows in a radial direction away from the so-called "focus of expansion" at a velocity which increases with its nearness to the observer and with its angular deviation from the observer's direction of gaze. Theoretically, the velocity field associated with such an optical flow pattern might provide visually effective information about both the egocentric distance of a given point - about its "time to contact", as Lee (1974) pointed out - and about the observer's direction of locomotion. Thus, the velocity fields associated with such optical divergence patterns might provide information about the relative depths of points on the surfaces of environmental objects.

When the direction of motion and the direction of gaze do not coincide or when the observer changes the direction of gaze during locomotion, the geometrical relations between the velocity field in the image and the distances of points form the observer become more complicated. The retinal image trajectories continue to point toward the vanishing point (the retinal image of the direction of gaze) despite changes in the relative direction of locomotion (see Regan & Beverley, 1982). The velocity field, however, is influenced by the direction of locomotion as well as by the distances of points from the observer. In principle, therefore, the velocity field might provide information about the orientation of a surface relative to the observer, as Prazdny (1983) and Perrone (1989) have indicated. This potential optical information about the structure and orientation of the surface is provided by the spatial derivatives of the velocities rather than by the directions of the image trajectories of the moving points.

So far as I am aware, however, human sensitivity to the differential structure of the velocity fields of these optical divergence patterns has not been shown to be sufficient for discriminating environmental surface structure. Indeed, experiments begun this summer at NASA-Ames by the working group on "Perceiving structure from motion" indicate that human sensitivities to this form of optical information are quite poor.

Observers were asked to discriminate the amount of slant of densely dotted planar surfaces away from the frontal parallel plane when these surfaces were displayed as if seen during translational motion in the direction of gaze - i.e., perpendicular to the display screen and moving toward the surface in question. As an additional visual reference, a ground plane was also visible, parallel to the simulated direction of motion and attached to the slanted surface along a horizontal line in the image.

All the observers of these displays were strikingly insensitive to the orientation of the simulated surface. The angle between the slanted plane and the frontal parallel plane was consistently and grossly underestimated, often by more than 45°. Even when the plane was nearly perpendicular to the direction of gaze, it often seemed slanted toward the observer. Moreover, the observers had little confidence in their judgments of the surface slant, and their judgments were inconsistent. Although we did not determine whether the observers might have been sensitive to the curvature of surfaces portrayed in this way by optical divergence patterns, the insensitivity to the angle between the ground plane and the slanted plane indicated that variations in curvature would not have been very visible. The perceived structure and orientation of the surface seemed to be influenced mainly by the 2D orientations of the image trajectories of the points in the optic flow pattern rather than by the velocity field as such, and these orientations are very poorly correlated with the relative depth or orientation of the surface.

49

In summary, presently available psychophysical evidence suggests that the divergence of optic flow fields is a poor source of information about environmental surface structure.

Next, we consider the perception of structure from motion parallax produced by translations in directions approximately perpendicular to the direction of gaze. Three aspects of the perception of these optical transformations are noteworthy: (1) These optical transformations produce much more accurate perception of environmental surface structure than the divergence patterns produced by translations parallel to the direction of gaze. (2) These transformations typically appear to have been produced by rotation rather than translation. (3) The tendency to perceive motion parallax patterns as rotation suggests how these perceptions are derived from local retinal information.

One of the questions examined in the experiment begun this summer at NASA-Ames was whether the perception of surface slant would be different when the simulated direction of translation was parallel to the slanted surface, so that the surface flowed horizontally over the display screen without changing the distance between the image and the surface. The motion parallax in these displays consisted of differential horizontal velocities which varied in inverse proportion to the distance of any given point from the observer's focal point-in contrast to the optical divergence patterns produced by moving in the direction of gaze toward the surface. The perceptual consequence of this change in the relative direction of motion was a dramatic improvement in the discriminability of the surface slant. This task was almost trivially easy in comparison with that when the direction of movement was toward the surface. Evidently, the optical information about surface structure was visually much more effective when the viewing position moved parallel to the surface.

Even though the optical transformation in the latter case was produced by a translation, the surfaces in these displays appeared to be rotating, as if the observer were moving around the arc of a large circle centered at some distant point beyond the field of view in the direction of gaze. This subjective impression is consistent with the theoretical idea that the perception of structure from motion is based mainly on optical transformations that are visually represented as rotations of surfaces in depth. Similar impressions of illusory rotation in motion parallax displays have also been observed by M. Braunstein and G. Andersen (personal communications, 1989).

Essentially the same phenomenon is involved in the "stereokinetic effect" (cf. Proffitt, Schmuckler, & Rock, 1989): In the standard demonstration, circular contours are arranged concentrically, centered about points that are laterally displaced from one another along a common invisible line. When these contours are rotated in the frontal parallel plane around a point at the center of the largest circle, the result is a strikingly compelling illusion of depth, with the contours whose centers are farthest from the center of the planar rotation appearing at the greatest depth from the plane of rotation and closest to the observer. (The Exploratorium in San Francisco has several fascinating demonstrations of this illusion.) This illusion results in part from local ambiguities about the direction of motion of the rotating circular contours, where the momentary velocities can be locally described by translations with a significant visible component perpendicular to the contour. Thus, differential local velocities are produced by the series of contours with varying curvatures and distances from the true center of rotation. The perceptual result is that the spatial pattern appears to be rigidly connected in depth and rotating around an axis which is tilted in depth in changing directions rotating around the line of sight. (More direct experimental evidence for this interpretation of the stereokinetic phenomenon will be reported by the author and his colleagues in the future).

50

The illusory rotation in depth that frequently occurs in these motion parallax patterns suggests some basic characteristics of the visual perception of spatial structure in moving optical patterns: First, the optical information that yields these perceptions seems to be local rather than global. Although rotations and translations may produce optical transformations that are globally quite different, even large global deformations that should in principle accompany the misperception of translations as rotations seem to go unnoticed; patterns which should appear plastic appear instead rigid. Local relations seem to govern the perceived global structure. These local relations are probably associated with spatial relations on connected surfaces.

Second, rotations seem to play a predominant role in visual representations of the optical transformation produced by moving objects and observers. The visual efficacy of these rotational representations may derive from the fact that translations may be locally approximated as rotations; the local first- and second-order derivatives are essentially the same in the two cases. The primacy of the rotational representations may be associated with preservation of local metric structure of a surface patch. Translations on the other hand, usually would not produce significant changes in the local differential structure of the image of a surface patch.

PERCEIVING THE 3-DIMENSIONAL FRAMEWORK OF ENVIRONMENTAL SPACE

So far, we have only examined the visual information about the surface structure of a single environmental object. This class of optical information does not specify the 3D structure of the space that contains that object; it does not specify the orientation of the object relative to either the observer or to some external reference; it does not specify the distance of the object from either the observer or from other separate objects; and it does not specify the location of the observer within this environmental space.

All of the latter properties involve the perspective optical projection from 3D Euclidean space (E^3) onto the 2D image surface (R^2). For any given local surface patch, this projective mapping is locally described by the six parameters of the 3×2 Jacobian matrix P, which embeds the retinal image of the surface patch into a specific coordinate system for E^3. As shown above, the values of these six parameters are not determined by the image transformations associated with rotation of the object; only the metric structure of the surface patch, described by the three metric tensor parameters of the matrix P*, can be derived from the invariance of object's structure under rotation.

The perspective projection from E^3 onto the retinal image is determined by the position of the observer's retina within the environment and by the direction of gaze. Thus, six parameters are needed to specify this perspective projection—three to specify the 3D position of the eye's focal point and three more to specify the fixation point or direction of gaze. These perspective parameters reflect global constraints among the local metric tensor parameters, P*, which embed images of the local surface patches into E^3. Similarly, the global perspective parameters also constrain the values of the local metric tensor parameters.

The perspective projective mapping onto the retinal image, from E^3 onto R^2, induces a version of hyperbolic geometry in the image: An infinite number of parallel lines may intersect at any given

51

point in the image. The lines which are regarded as parallel in the hyperbolic geometry of the 2D retinal image are the perspective images of lines that are parallel in E^3. The retinal images of lines that are parallel in E^3 converge at a point in the retinal image which is the image of an environmental point infinitely distant from the observer in that direction. Thus, if the observer were standing in a flat open field with no changes in elevation and looking "straight ahead" in a direction parallel to the ground plane, the images of parallel lines extending into the distance parallel to the direction of gaze would converge at a vanishing point on the horizon that is sometimes called the "center of vision." Other sets of parallel lines that extend in other directions parallel to this same ground plane will also converge at other image points that lie along the same horizon line.

This horizon line is important in the geometry of vision because it represents the observer's eye-height: The images of environmental objects above the observer's eye- height lie above the horizon line, and the images of objects below the observer's eye- height lie below the horizon line in the retinal image. Thus, the horizon line divides the retinal image into two regions, one region above and the other below the observer's eye. In most visual environments, however, the horizon line is not explicitly visible. But even when it is not explicitly visible in the retinal image it is implicitly specified by the convergence of image lines that are parallel to the ground plane. Thus, for example, if an observer is standing inside a rectangular room, the four lines defined by the intersections of the side walls with the floor and ceiling project onto the retinal image as four lines which if extended would cross at a single point corresponding to the observer's eye-height.[4]

Now it is useful to consider the retinal surface as a section of a sphere centered at the focal point of the observer's eye. This spherical set of potential visual images constitutes what is known as the optic array (cf. Gibson, 1966; Ch. 10; Cutting, 1986, Ch. 2; Johansson & Börjesson, 1989). Thus, the horizon line extended in all directions a full 360 degrees around the observer would define a great circle in the optic array—the intersection of the sphere with a plane passing through its center parallel to the ground plane, dividing the optic array into two equal hemispherical sections, one containing the images of objects above and the other containing the images of objects below the observer's eye. The shapes and locations of the images in the optic array are determined by the shapes and locations of environmental objects and by the location of the observer's station point within the environment. By definition, the optic array remains invariant under rotations of the eye, though of course the retinal positions of the images of objects are altered as the observer rotates his or her eye to look at various environmental objects.

The spatial relations associated with the position of the retina within the optic array constitute an important scourge of optical information about the orientation of the eye within the environment. Thus, pitch (rotation around the horizontal axis) is described by the elevation of the horizon line in the retinal image; roll (rotation around the "depth" axis parallel to the ground plane) is described by the angular orientation of the horizon line in the retinal images; and yaw (rotation around the vertical axis perpendicular to the ground plane) affects only lateral translation in the retinal image.

The importance of such spatial information for the observer's perceiving his or her orientation within the environment has been demonstrated in several recent psychophysical studies by Johansson and Börjesson (1989), Matin and Fox (1989), and Stoper and Cohen (1989). Changes in the orientation of a structured optical pattern were shown in these studies to exert a large influence on the perceived orientation of the (gravitational) ground plane. Although an explicit horizon line

was not visible in these studies, its location was implied by the convergence of straight lines which appeared to be parallel with each other and with the ground plane. Such optical information associated with the vanishing points of parallel lines is perceptually compelling, capable of dominating contradictory knowledge and proprioceptive information about the direction of gravity.

Although the horizon line is an important aspect of the geometry of vision, its visibility should not be overemphasized: The position of the horizon line in the retinal image is no more visible than the orientation of the ground plane. The horizon line is simply the locus of vanishing points of lines parallel to the ground plane. If one does not know which lines are parallel to the ground plane, then neither does one know the location of the horizon line. Moreover, the horizon line is not a unique structural characteristic of the visual field; any point in the visual field can be the vanishing point of parallel lines in that direction.

Two lines parallel in E^3 but not parallel to the ground plane will converge in the image at a point that does not lie on the horizon line. Thus, for example, sets of parallel lines that are parallel with a vertical plane which is perpendicular to the horizon line would converge in the optic array at a point on a great circle which is perpendicular to the horizon line. If such a vertical arc passes through the center of vision, it divides the visual field into left and right halves, separating objects which lie to the left and right of the observer's direction of gaze. The full set of great circles passing through the center of vision forms a polar configuration radiating from the center of vision. Of course the polar configuration defined by this set of great circles is not unique to the center of vision. Similar sets of great circles can be described at any given point in the image. Every great circle in the optic array is the image of points that are infinitely distant from the observer in some direction parallel to a plane that contains the observer's station point.

Any set of parallel lines in E^3 is parallel to two orthogonal planes through the station point, and they would converge in the image at a point that is an intersection of the corresponding two orthogonal great circles in the optic array. Three parameters are needed to specify each of these vanishing points on the optic array—two parameters to specify any point on the sphere and another parameter to specify the orientation of the two orthogonal great circles at that position.

The physical significance of these vanishing points may be appreciated by considering the images of objects translating through the environment in a constant direction, with no rotation, as seen by an immobile eye of a stationary observer. The trajectories of all points on the object are parallel in E^3 and the images of these trajectories are straight lines which would converge in the image at a specific vanishing point. These converging image lines produced by an object's linear trajectory are also parallel in the hyperbolic geometry of the optic array.

Let us now consider a subset of these images of moving objects—those whose linear trajectories are normal to the spherical optic array, passing through its center at the observer's station point. The stationary images of such a moving object remain congruent with each other in the hyperbolic geometry of the image (even though their size changes in the Euclidean geometry of the image). Congruence is a property possessed by hyperbolic geometry as well as by Euclidean geometry. This congruence of the images of objects under this class of translational motions in E^3 serves to specify the structure of the 3D space in which these motions and objects occur. Figures 2 and 3 illustrate how such hyperbolic isometries in the image specify Euclidean isometries in an environmental 3-space. In

53

both of these illustrations the structure of the space is described by the congruence or symmetry of stationary objects repeated at varying locations throughout the space and its image. In natural visual environments such isometries are often revealed temporally, by the sequential images yielded by moving objects and moving observers.

When the trajectory of an object moving in relation to the observer is in a direction that does not coincide with the focal point of the eye, then its sequential images are not strictly congruent with one another in the hyperbolic geometry of the visual image sphere. In addition to the changes in size (in the Euclidean sense) produced by the changing distance between the object and the eye, the projected shape of any given surface patch also undergoes a deformation associated with a relative rotation in E^3, as described in the preceding sections of this paper. Such deformations of projected shape may be seen in vertical or diagonal sets of images of the square forms in Figure 2 or in corresponding directional sets of images of the flying-fish-like forms in Figure 3. Despite these projective deformations, the implied congruence of objects under motion in E^3 is immediately visible. That is, a fundamental characteristic of E^3 is its isometry under translations and rotations in any of the three orthogonal directions, and this isometry is displayed by the congruence of the sequential images of objects moving in relation to the observer.

The image transformations produced by the observer's translational motion through the environment are especially informative—about the scaling of the relative sizes of environmental objects, about the scaling of the relative distances of objects from the observer and from each other, and about the relative location and motion of the observer within the environment. These observer-produced image transformations are informative because they yield globally parallel trajectories in E^3 for the relative motions of points on environmental objects and surfaces - trajectories which diverge in the optic array from the horizon line and from the direction of locomotion, thereby describing and scaling the hyperbolic geometry of the optic array as an image of the environment.

Contrary to what might be supposed, these optical relations are more informative about the observer's direction of locomotion and involve disconnected points distributed over varying directions and distances from the observer. Cutting (1986) provides convincing experimental data on this effect, showing that the relative motions of contours lying directly ahead in a plane perpendicular to the path of locomotion yield much less accurate judgments about the direction of locomotion than do those laterally displaced from the path of locomotion. The parallel image trajectories of discrete texture elements distributed over an extended ground plane have been shown to provide sufficient optical information for accurate judgments of the direction of locomotion along linear (Warren, Morris, & Kalish, 1988; Warren & Hannon, 1988) and even curvilinear paths (Warren et al., in press). Recent results of G. J. Andersen (personal communications; Andersen & Dyre, 1989) indicate that similar performance can also be obtained from patterns of discrete points randomly distributed in a 3D cloud-like volume displayed as if the observer were translating through the cloud. Evidently, the global hyperbolic geometry of the optic array and hence the E^3 geometry of the environmental layout are best revealed by the divergence component common to the trajectories of spatially separate contours and edges distributed throughout the visual field.

The present geometric analysis of the optical information for perceiving one's own position and motion within the environment differs in two noteworthy respects from many other contemporary analyses: First, the hyperbolic geometry of the perspective projection of the environmental E^3 space

has been described in terms of the spherical optic array rather than the retinal images. By definition, the optic array remains invariant under rotational eye movements. In contrast, the velocities and directional trajectories of the retinal images of moving environmental objects are significantly affected by the eye movements involved in fixating or tracking various environmental objects. Accordingly, the changing optical patterns associated with the observer's motion through the environment constitute a much less direct source of information about the observer's position and motion in the environment when the spatial structure of these optical patterns is described by reference to the retinal coordinates. As emphasized in the earlier sections of this paper, however, there seems to be no compelling empirical or theoretical requirement for assuming that the spatial relations detected by vision must be referenced to the local retinal coordinates rather than to the neighboring optical pattern. In any case, the present analysis is based on the optic array, in which the spatial structure remains invariant under rotational eye movements.

Second, the present analysis is based on the spatial structure of the optic array and the transformations of this structure produced by moving objects and observers. In contrast, many contemporary theoretical analyses of optic flow have focussed on the velocity field (e.g., Cutting, 1986; Prazdny, 1983). In the present analysis, visible spatial information is provided by the spatial structure associated with parallelism of the image trajectories and with congruence of the successive images of moving objects. That is, if the moving optical patterns are represented as vector fields, where the velocity of a given point is represented by the length of an associated vector, then the visually detected spatial information is assumed to be associated with the directions rather than the lengths of these vectors.

Finally, we note again that the global nature of this optical information about the hyperbolic geometry of the spatial layout of the environment and the observer's position within it contrasts with the local nature of the optical information about the smooth surface structure of a single object. The former global information derives from parallelism and congruence associated with translations, whereas the latter information derives from local deformations produced by rotations. Presumably, the visual mechanisms for detecting these two functionally different classes of geometric information also differ from one another.

REFERENCES

Andersen, G. J. & Dyre, B. P. (1989). Spatial orientation from optic flow in the central visual field. Perception and Psychophysics, 45, 453-458.

Cutting, J. E. (1986). Perception with an eye for motion. Cambridge, MA: MIT Press, Bradford Books.

Doner, J., Lappin, J. S., & Perfetto, G. (1984). Minimal conditions for the visual detection of structure and motion in three dimensions. Science, 209, 717-719.

Gibson, J. J. (1966). The senses considered as perceptual systems. Boston: Houghton-Mifflin.

Guggenheimer, H. W. (1963). Differential geometry. New York: McGraw-Hill. (Reprinted by Dover Publications, New York, 1977).

Johansson, G. & Bîrjesson, E. (1989). Toward a new theory of vision: Studies in wide- angle space perception. Ecological Psychology, 1, 301-331.

Koenderink, J. J. & van Doorn, A. J. (1976). The singularities of the visual mapping. Biological Cybernetics, 24, 51-59.

Lappin, J. S. (in press). Perceiving the metric structure of environmental objects from motion, self-motion and stereopsis. In R. Warren & A. H. Wertheim (Eds.), The perception and control of self-motion. Hillsdale, NJ: Erlbaum.

Lappin, J. S., Doner, J. F., & Kottas, B. L. (1980). Minimal conditions for the visual detection of structure and motion in three dimensions. Science, 209, 717-719.

Lappin, J. S. & Fuqua, M. A. (1983). Accurate visual measurement of three-dimensional moving patterns. Science, 221, 480-482.

Lappin, J. S. & Love, S. R. (in preparation). Visual measures of stereoscopic form from congruence under motion.

Lee, D. N. (1974). Visual information during locomotion. In R. B. MacLeod & H. L. Pick, (Eds.), Perception: Essays in honor of J. J. Gibson. Ithaca, NY: Cornell University Press.

Matin, L. & Fox, C. R. (1989). Visually perceived eye level and perceived elevation of objects: Linearly additive influences from visual field pitch and gravity. Vision Research, 29, 315-324.

Perrone, J. A. (1989). In search of the elusive flow field. Proceedings of the workshop on visual motion. Washington, DC: Computer Society of the IEEE.

Prazdny, K. (1983). On the information in optical flows. Computer vision, graphics, and image processing, 22, 239-259.

Proffitt, D. R., Rock, I., & Schmuckler, M. (1989). Visual bases of the stereokinetic effect. Paper presented at the Fifth International Conference on Event Perception and Action, Miami University. Oxford, OH, July, 1989). Manuscript n preparation.

Rogan, D. & Beverley, K. I. (1982). How do we avoid confounding the direction we are looking and the direction we are moving? Science, 215, 194-196.

Stoper, A. E. & Cohen, M. M. (1989). Effect of structured visual environments on apparent eye level. Perception and Psychophysics, 46, 469-475.

Warren, W. H., Jr. & Hannon, D. J. (1988). Direction of self-motion is perceived from optic flow. Nature, 336, 162-163.

Warren, W. H., Jr., Mestre, D. R., Blackwell, A.W., & Morris, M. W. (in press). Perception of curvilinear leading from optical flow. Journal of experimental psychology: Human perception and performance.

Warren, W. H., Jr., Morris, M.W., & Kalish, M. (1988). Perception of translational heading from optical flow. Journal of experimental psychology: Human perception and performance, 14, 646-660.

NOTES

1. Preparation of this report was supported in part by NIH Grants EY-05926 and P30- EY-08126. The mathematical ideas have been greatly influenced by John Ratcliffe (Dept. of Mathematics, Vanderbilt University) and by discussions with Alan Peters (Dept. of Electrical Engineering, Vanderbilt University).

2. Three technical qualifications bear mention: First, this isomorphism applies, of course, only to the visible regions of the surface. On any given curved surface, especially those surrounding opaque solid objects, some source regions of the surface will generally be occluded from view by other regions of the same surface or by other separate surfaces which are closer to the observer in the same visual direction. Second, this isomorphism also assumes that the scale of resolution with which the environmental surface is described corresponds with that of its image and that this scale of resolution falls within the range of resolution capabilities of the visual system. Third, some surfaces are transparent, resulting in the images of separate surfaces superimposed on the same retinal location. Nevertheless, none of these three technical qualifications should be considered to invalidate the essential correspondence between the two manifolds.

3. The term metric is used in the conventional mathematical sense: A relation $m(a,b)$ between two elements a and b is said to be a metric relation if it satisfies the following axons for all a, b, and c: (i) non-negativity: $m(a,b) \ Ú \ 0$; (ii) symmetry: $m(a,b) = m(b,a)$; (iii) reflexivity: $m(a,a,) = 0$; (iv) triangle inequality: $m(a,c,) \ Û \ m \ (a,b) + m(b,c)$. Euclidean distances in abstract empty space constitute a special case of metric relations. Of particular interest in the present context are metric relations over curved surfaces, which remain invariant under bending of the surface.

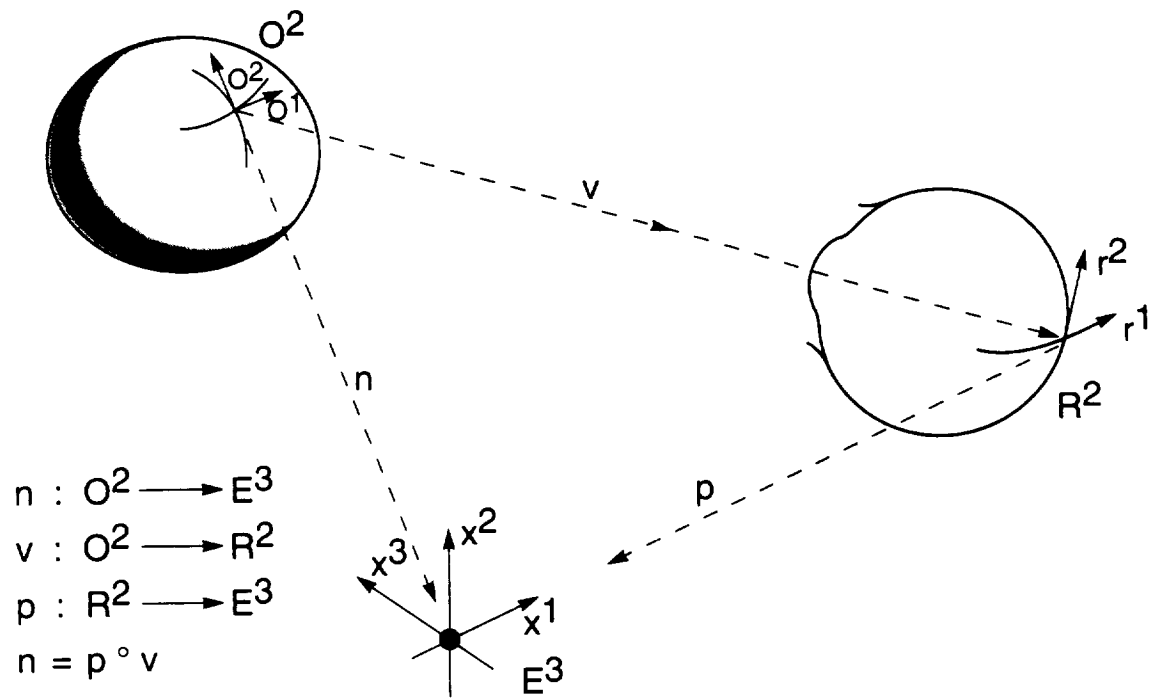4. I am grateful to Steven Tschantz for pointing this out to me.

Figure 1. A schematic illustration of the relationship between three separate coordinate systems for describing the surface structure of an environmental object and its image, and the mappings between these coordinate systems.
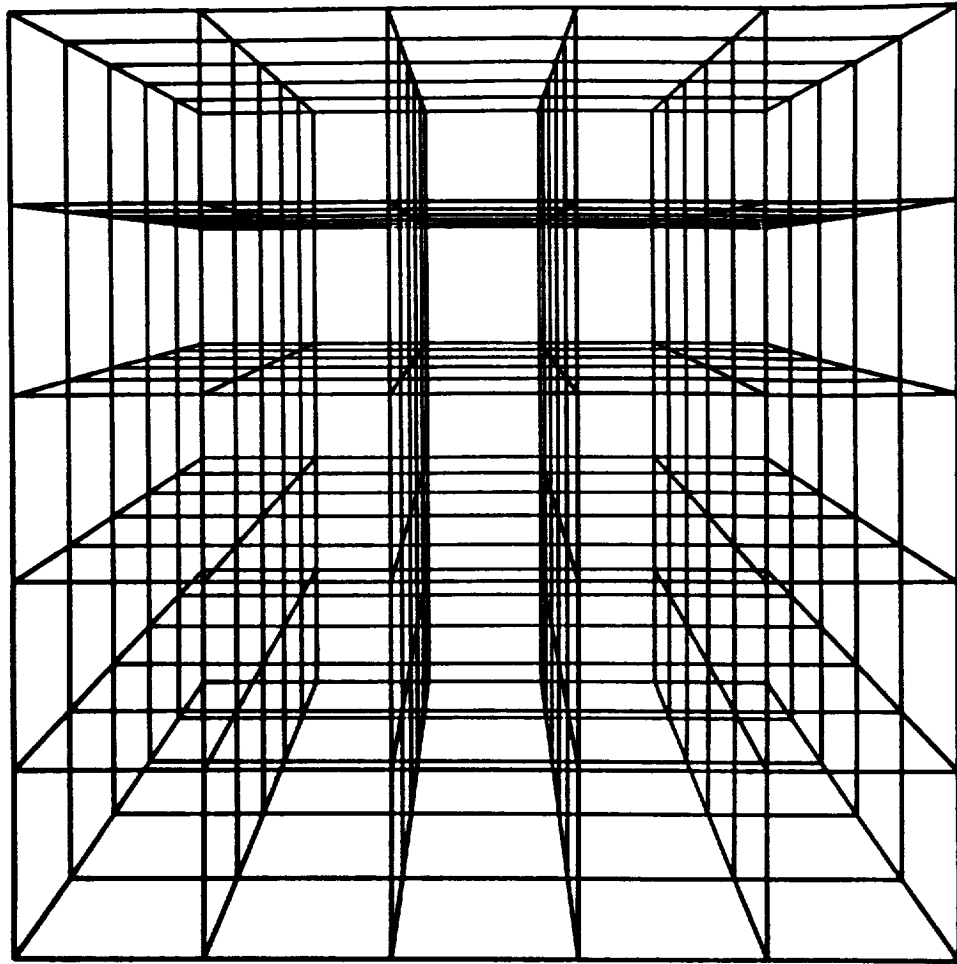
Figure 2. A perspective image of a 5 × 5 cube. (I am indebted to Steven Tschantz for providing this illustration.)

Figure 3. Depth - a wood engraving by M. C. Esher, 1955.

THE PERCEPTION OF SURFACE LAYOUT DURING LOW LEVEL FLIGHT

John A. Perrone
NASA Ames Research Center
Moffett Field, California and
Stanford University
Stanford, California

INTRODUCTION

Low level flight in helicopters presents a particularly challenging visual situation to the pilot, especially if his normal field of view is degraded or restricted by the use of night-vision aids. The pilot must somehow extract information about his heading direction and the layout of the environment in front of him from the two-dimensional pattern of light reaching his retina. Humans can perform this task well in most situations, but low-level flight taxes this ability to the limit, because the time to respond to obstacles is greatly shortened and there is often a large rotational component to the observer platform that is not normally present in normal ambulatory motion.

There are many sources of information that a pilot can use to infer the three dimensional environmental layout from the two-dimensional images that feed into his visual system. This paper will concentrate on one of these sources, namely the 2-D motion flow-field that is generated during observer motion. We are interested in how the pilot can infer his heading and the surface layout from these 2-D velocity flow fields. This is part of the more general "structure from motion problem" and the special case of observer motion has been labelled the structure from ego-motion problem.

There now exists many theoretical treatments of this problem, which demonstrate how 3-D information can be extracted from the two-dimensional motion field. (e.g. [1, 2, 3, 4, 5]). However, explicit theories as to how human observers actually extract this information are not so common. Many of the above theories also assume that the 2-D flow field is available, but this stage can represent a major stumbling block in trying to solve the structure from motion problem [6].

This problem is obviously very relevant to the visually guided control of movement. It is a very difficult problem, however, and one which may not be fully understood for some time. A complete understanding of the process would help detect potentially dangerous situations (illusions) and help in the design of display systems such as FLIR. The designers could economize on features that were known to be unimportant to the control problem and make more informed decisions about parameters such as the field of view and the amount of visual noise that can be tolerated in these displays.

Obviously visual motion is an important source of information for man and many animals [7, 8, 9, 10]. However, in the context of rotorcraft flight the question remains as to just how much information about surface layout and obstacles can be derived from motion cues ? We know that heading information can be gained from the motion field [7, 11] and that relative motion and parallax can

provide important information about depth discontinuities [1, 12]. What is not so well established concerns the case of an observer moving forward with a fixed line of sight. This is supposedly the simplest situation for the structure from motion problem (since there is no rotational component), yet it not obvious in this case as to how well the environmental layout can be inferred on the basis of motion information alone. If a pilot is viewing FLIR imagery of a very unstructured scene (e.g. foliage or trees) such that the usual cues of linear perspective, intensity gradients and size gradients are missing, how well can they perceive the scene layout on the basis of motion alone ? It is a useful exercise to examine some of the issues involved in this relatively simple situation. From this we can begin to see some of the important variables that need to be considered and which factors are worth manipulating experimentally. It also helps define the scope of the problem and puts a cap on what can realistically be obtained from machine vision applications such as remote sensors for autonomous flight. Identifying potential problem areas for computer vision systems can provide useful insights into what may also prove problematical to humans.

EFFECTIVE RANGE OF MOTION INFORMATION

One point that is often ignored is that the effective range over which motion information is even detectable is quite limited. There has been some work on this area in relation to the extraction of heading direction [11, 13] but this hasn't been carried over much to the structure from motion problem. Figure 1 shows the theoretical flow field for a craft moving at 3 eyeheights/sec over a flat plane. The figure shows a .25 sec "snapshot" of the motion field. Each vector is for a point on a square grid with the inter-point distance equal to 1 eyeheight. The thing to notice is that the length of the velocity vectors fall off rapidly with distance. If for simplicity we limit ourselves to points lying along the median plane, the angular velocity of the points can be found from:

$$\dot{EL} = -\frac{\dot{x}}{z}\sin^2 el \qquad (1)$$

where is the forward velocity parallel to the ground plane, z is the height of the observer above the ground plane, and el is the elevation angle (measured from the horizon) of the point on the ground. (Using equation 22 from Warren [14] with = 0). Substituting z/ for sin(el) we have:

$$\dot{EL} = -\dot{x}\frac{z}{x^2 + z^2} \qquad (2)$$

This shows that angular velocity basically falls off as the square of the distance from the observer. The angular velocity is also a function of the height of the observer, but the effect of the z term is less than the distance factor. Figure 2 shows a plot of angular velocity (in min of arc/sec) against distance (in height units) for an eyepoint moving at 3 eye-heights/sec. Points higher up in the field (smaller z) will have an even smaller value of and would lie below the curve shown in the figure. It is difficult to set a threshold level for velocity detection and it is based on many factors [15]. However, if we use the value based on practical complex situations [16] i.e 40 min of arc/sec we find that absolute velocity information is becoming subthreshold at about 15 to 16 eyeheight units away. This

64

is the length of the "headlight beam" defined by motion information alone. At a speed of 3 eyeheight units/sec, this only gives about 5 seconds to respond to features on the ground that are revealed by the motion process. These critical distances and times are shorter for objects above the ground that lie closer to the eye-level plane.

This analysis is only supposed to provide a qualitative feeling for the limitations to the structure from egomotion problem. It ignores the issue of how the structure is actually extracted and is based on absolute motion thresholds rather than the perhaps more relevant relative motion thresholds. However it does help to limit the domain over which motion information can be considered to contribute to the perception of surface layout. It could also be an important factor when the visual motion information is being relayed to the pilot via some form of CRT display device, such as those used in FLIR systems. The limited resolution of these displays degrades the motion information even further and thus puts a further cap on the effective range of the structure from motion process.

This limited range of utility for motion information also becomes an important issue when certain environmental features are present in the field of view. Most terrain is not perfectly flat like the example shown in Figure 1. Rather NOE flight is often over sloping terrain with many hills and valleys. The correct perception of the slope and orientation of this terrain is important to the pilots perception of his own spatial orientation. Any misperception of surface layout can affect the flight path chosen by the pilot and, in theory, could lead to disorientation in some cases.

It is therefore of interest to examine the motion information that is available during flight towards sloped terrain.

Motion Information and Sloped Terrain

It turns out that this situation generates an interesting pattern of flow information and raises important theoretical issues as to how humans actually infer layout from the motion flow field. Figure 3 shows the theoretical flow field for pure translatory motion towards a planar hill, slanted relative to a ground plane. The angular velocity of points on the hill decrease as a function of distance from the observer as before, but also as a function of height in the field. This is the z term in the numerator and denominator of equation 2 above. The z term decreases as a function of the distance as well as a function of the tangent of the slant angle of the hill. This means that for situations such as that shown in Figure 3, the angular velocity is low over a large area of the field, especially in the important region directly ahead of the observer. This makes it difficult for any system attempting to infer the slope of the terrain using the angular velocity of points in the field.

There are many ways by which the slant could be recovered from the motion information. It is interesting to consider some of the techniques that an artificial vision system could use to tackle this problem. One technique would be to assign a depth value to each of the points based on its angular velocity (i.e. the length of the vectors in fig 3). Since we assume that the actual observer speed is unknown, this amounts to the derivation of a relative range map, or a time to impact map which is independent of the actual speed [17]. The slant would then be found by "fitting" a plane to the distribution of relative distances derived from the impact times. However in order to convert the angular velocities into impact times, the magnitude of each velocity vector must be divided by a factor proportional to the square of the distance in the image plane, of the point from the focus of expansion.

Because the angular velocities are so small for a large proportion of the points in the field, we cannot expect the impact time map to be very accurate in these regions, especially when the velocity estimates are noisy.

The other general approach would be to use relative velocities (e.g. [1]) and find the slant of a local patch and then integrate over the entire surface. However since this requires taking differences of already small 2-D velocity estimates, we can again expect errors to occur over a large part of the field. The problem is that the relative difference in velocities in small local patches can be quite small, and it is only when the differences in velocities of very disparate points are considered, that such differences may reach threshold. Techniques which use local differencing methods are good for finding depth discontinuities in the field, but with slanted surfaces the change can be very gradual.

There is an interesting parallel between the problem of judging slant from motion with the problem of detecting slant using stereoscopic vision. Gillam et al., [18] have tried to argue that stereoscopic process requires changes (2nd derivatives) in the slant to perceive the slant correctly. They found that subjects took very long times to accurately judge the slant of flat planes, but much shorter times if the plane contained changes in slant or a discontinuity in depth. If it can be shown that the structure from motion system is also dependent upon such changes, then we can expect problems in situations where they are absent or subthreshold. Fortunately, most natural scenes and terrain are not perfectly flat like figure 3. There are instances however, such as with snow covered terrain, where many of the small details and features which can provide information about changes in slant are lacking. In such situations, the unique pattern of velocity vectors produced by slanted terrain during forward translation may become important. Relatively small differences in speed are produced locally and this could result in problems for a system that is designed for the detection of the large local speed differences. Such large differences are much more common in the visual field since they result whenever we move through and environment made up of objects occupying different depth planes.

There is more to the problem of detecting surface layout in the presence of slant however, than just the slow change in speed values. Braunstein, [19] showed that surface slant could be judged reasonable accurately for surfaces slanted only 20 to 30 deg from vertical. In this case, the local differences in speed are small. However Braunstein used motion parallel to the image plane so the the image motion was completely unidirectional. With forward motion, the velocity flow field consists of vectors of many different directions. Figure 4 shows what the flow field for the surfaces in figure 3 would look like in the case of translation parallel to the image plane. We are currently testing experimentally whether the difference in the flow fields between the two situations can account for any differences in the ability to extract surface slant under the two types of translation.

Comparison of the two flow fields in Figs 3 and 4 draws attention to another interesting feature. In the case of forward translation (fig 3) the perspective indicated by the vectors conflicts with the perspective indicating the static layout of the surfaces (the rectangular layout of the points helps to define the static perspective cues). The pattern defined by the motion vectors is similar to what would be obtained if a curved surface was covered with poles set normal to the surface. In the case of motion towards a flat plane (wall, cliff) (figure 5) the perspective suggested by the vectors is similar to that produced by a forward slanting surface like a ceiling.

If a display is used with a slow phosphor such that streaking of the image is present, one would predict some misperceptions of the surface layout to occur under these situations. We have noticed exactly these effects in our experimental displays when the image contrast is high and the room illumination low. It would be of interest to see if any of these effects can occur in cockpit mounted displays or FLIR displays. This is an instance of motion interacting with the pattern processing system. Motion artifacts such as streaking seem to be creating a misperception of surface orientation. This introduces a slightly different question: Can motion override a familiar "illusion" of surface orientation that occurs under static viewing conditions?

Motion and Surface Slant Underestimation

Many situations have been identified in which surface orientation under static viewing conditions, can be greatly misperceived by observers. Some of the stimulus features that contribute to these errors have been identified [20, 21, 22, 23]. One particular situation relevant to N.O.E flight conditions is the perceived slope (slant) of the terrain immediately in front of the craft. Under both laboratory and environmental testing conditions the slant of a surface is, in most cases, perceived to be closer to the observers frontoparallel plane than its true position [20, 21, 23]. This has mainly been tested under monocular static conditions although there is some evidence that slant underestimation still occurs for stereoscopic displays [24].

One question that has not been addressed fully in the research literature in this field is whether or not motion information can override this tendency to misperceive the slant of surfaces. For stimulus displays in which the motion is parallel to the image plane, Braunstein, [19] has shown that motion information provides strong cues for slant and that very little underestimation occurs under the motion conditions. It is not clear whether or not this same state of affairs would exist for the case of motion towards the surface. We have begun a series of experiments aimed at answering this question.

Although it is fairly well established now that much information about surface layout can be gained from motion cues, it is not so clear as to what information humans can use and what specific information they should be provided with. The various theoretical analyses tell us that the information is there in the stimulus. It will take many more experiments to verify that this information can be used by humans to extract surface layout from the 2-D velocity flow field. Pilots obviously can use the information efficiently in most situations. This paper has tried to draw attention to some of the visual motion factors that can affect the pilot's ability to control his craft and to infer the layout of the terrain ahead of him.

REFERENCES

1. Clocksin, W. F. Perception of surface slant and edge labels from optical flow: a computational approach. 9: 253-269, 1980.

2. Longuet-Higgins, H. C. A Computer Algorithm for Reconstructing a Scene From Two Projections. 293: 133-135, 1981.

3. Prazdny, K. Determining the Instantaneous Direction of Motion From Optical Flow Generated by a Curvilinearly Moving Observer. : 109-114, 1981.

4. Tsai, R. Y. and T. S. Huang. Uniqueness and Estimation of 3-D Motion Parameters and Surface Structures of Rigid Objects. Computer Science. : 1-34, 1983.

5. Zacharias, G. L., A. K. Caglayan and J. B. Sinacori. A visual cuing model for terrain-following applications. J. Guidance. 8(2): 201-207, 1985.

6. Perrone, J. A. In search of the elusive flow field. Proceedings of the IEEE Workshop on Visual Motion. 181-188, 1989.

7. Gibson, J. J. "The perception of the visual world." 1950 Houghton Mifflin. Boston.

8. Nakayama, K. Biological image motion processing: a review. Vision Res. 25: 625-660, 1985.

9. Hildreth, E. C. and C. Koch. The Analysis of Visual Motion: From Computational Theory to Neuronal Mechanisms. Ann. Rev. Neurosci. 10: 477-533, 1987.

10. Horridge, G. A. The evolution of visual processing and the construction of seeing systems. Proc. R. Soc. Lond. 230: 279-292, 1987.

11. Cutting, J. E. "Perception with an eye for motion." 1986 Bradford. Cambridge.

12. Nakayama, K. and J. M. Loomis. Optical velocity patterns, velocity-sensitive neurons, and space perception: A hypothesis. Perception. 3: 63-80, 1974.

13. Palmer, E. A. Experimental Determination of Human Ability to Perceive Aircraft Aim Point From Expanding Gradient Cues. 40th Annual Scientific meeting of the AMA. 176-177, 1969.

14. Warren, R. Optical Transformation During Movement: Review of the Optical Concomitants of Egomotion. O.F.O.S.R. report 81-0108. 1982.

15. Salvatore, S. The Perception of Real Motion A Literature Review. Injury Control Research Laboratory Report ICRL-RR-70-7. 1972.

16. Michaels, R. M. and L. W. Cozan. Perceptual and field factors causing lateral displacements. Highway Research Record. 25: 1963.

17. Lee, D. N. "Visual information during locomotion." Perception: Essays in honor of J.J. Gibson. Macleod and Pick ed. 1974 Cornell U. Press. Ithaca.

18. Gillam, B., T. Flagg and D. Finlay. Evidence for disparity change as the primary stimulus for stereoscopic processing. Perception and Psychophysics. 36: 559-564, 1984.

19. Braunstein, M. L. Motion and texture as sources of slant information. J. Exp. Psychology. 78: 247-253, 1968.

20. Gibson, J. J. The Perception of Visual Surfaces. Am. J. of Psych. 63: 367-384, 1950.

21. Clark, W. C., A. H. Smith and A. Rabe. The Interaction of Surface Texture, Outline Gradient, and Ground in the Perception of Slant. Canad. J. Psychol. 10(1): 1-8, 1956.

22. Gogel, W. C. Equidistance Tendency and its Consequences. 64(3): 153-163, 1965.

23. Perrone, J. A. Visual slant underestimation: a general model. Perception. 11: 641-654, 1982.

24. Gillam, B. J. Perception of Slant when Perspective and Stereopsis Conflict: Experiments with Aniseikoniclenses. J. of Exp. 78: 299-305, 1968.
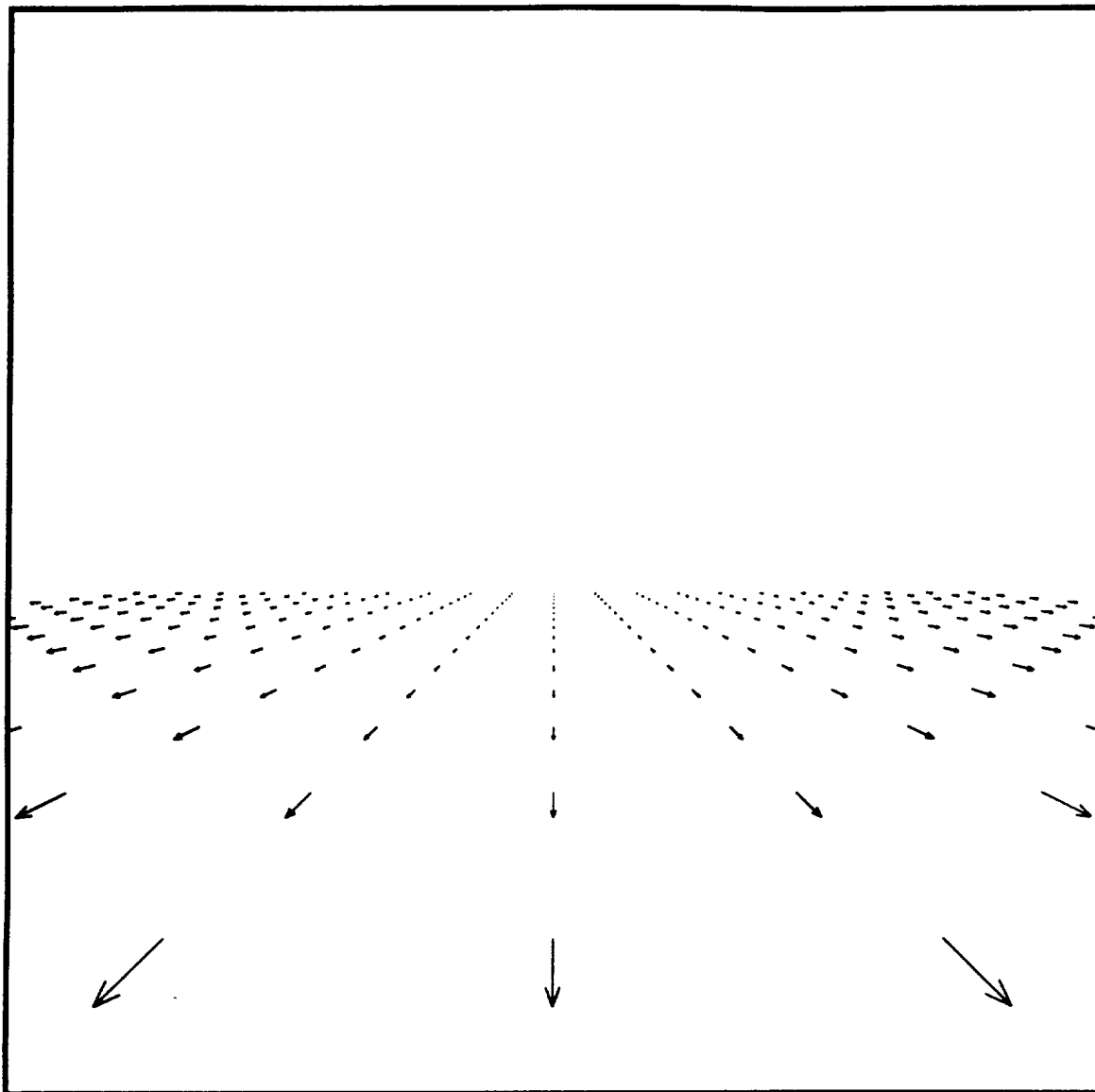
Figure 1. Theoretical flow field for motion over a flat ground plane.
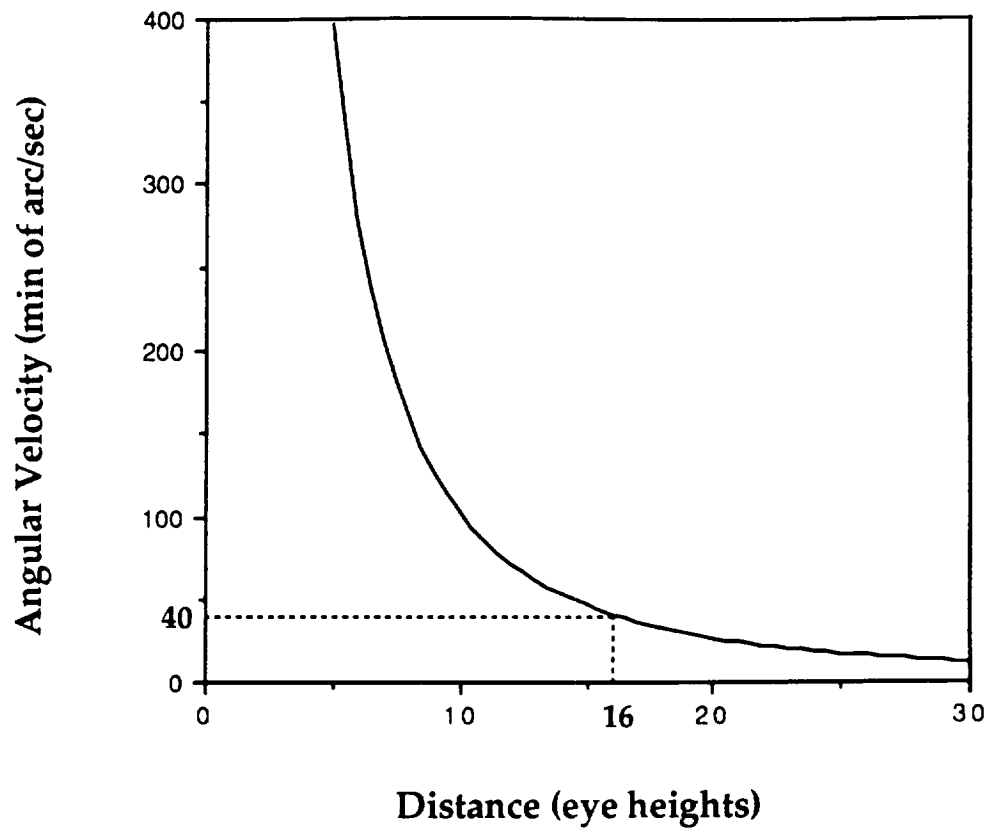
Figure 2. Plot of angular velocity versus distance along ground plane.
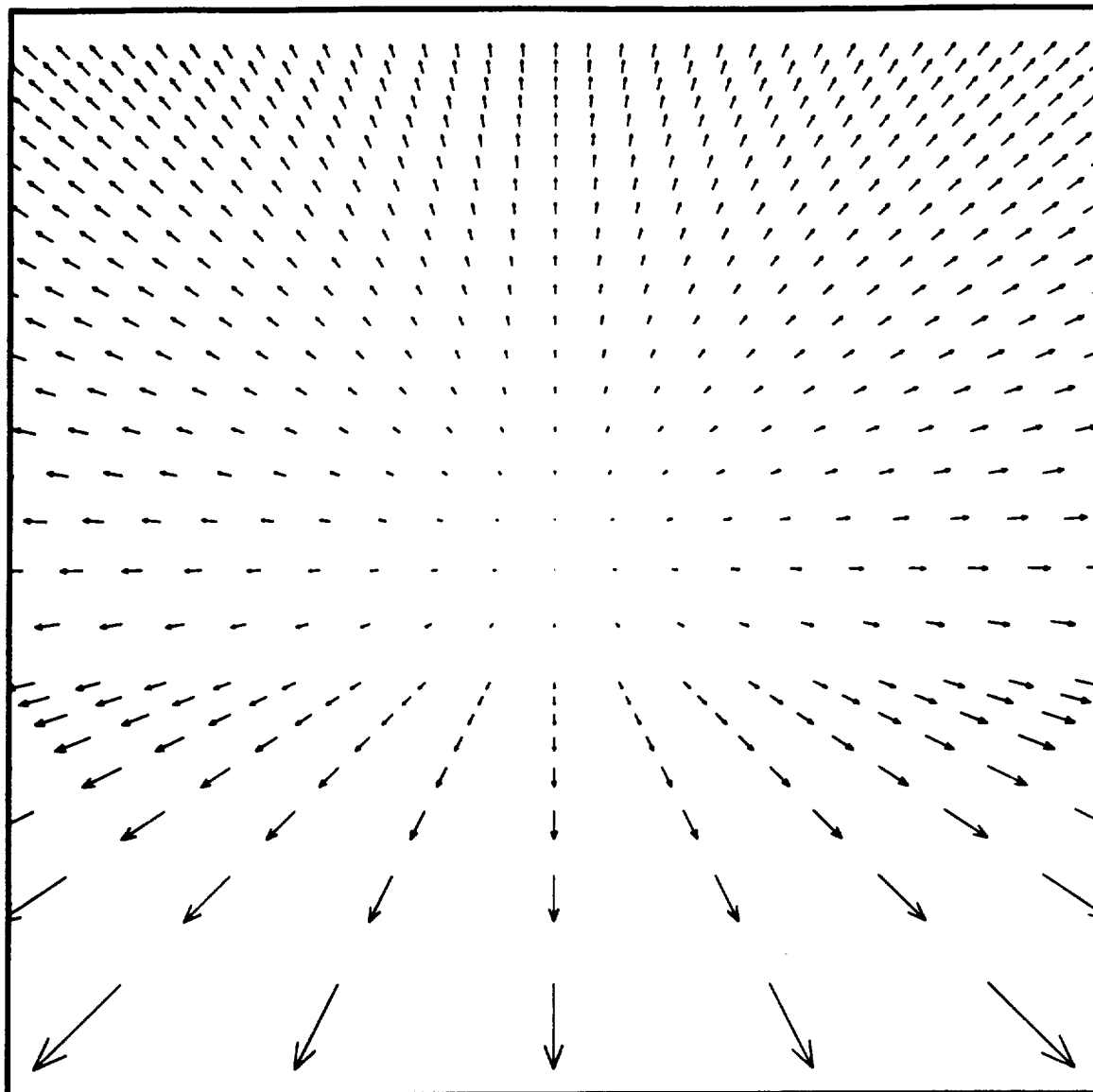
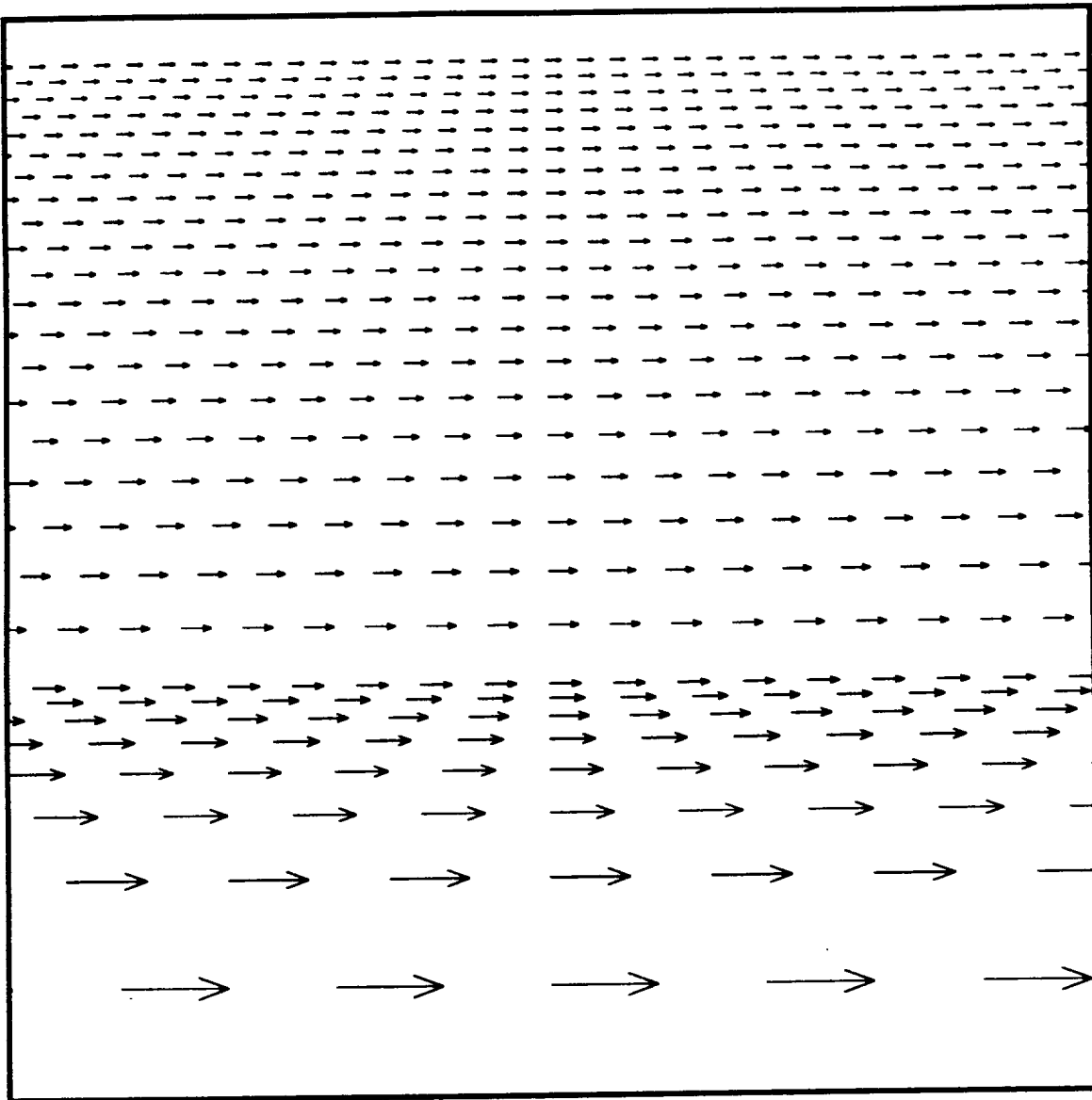Figure 3. Flow field for motion toward a plane slanted 60 degrees from the horizontal.

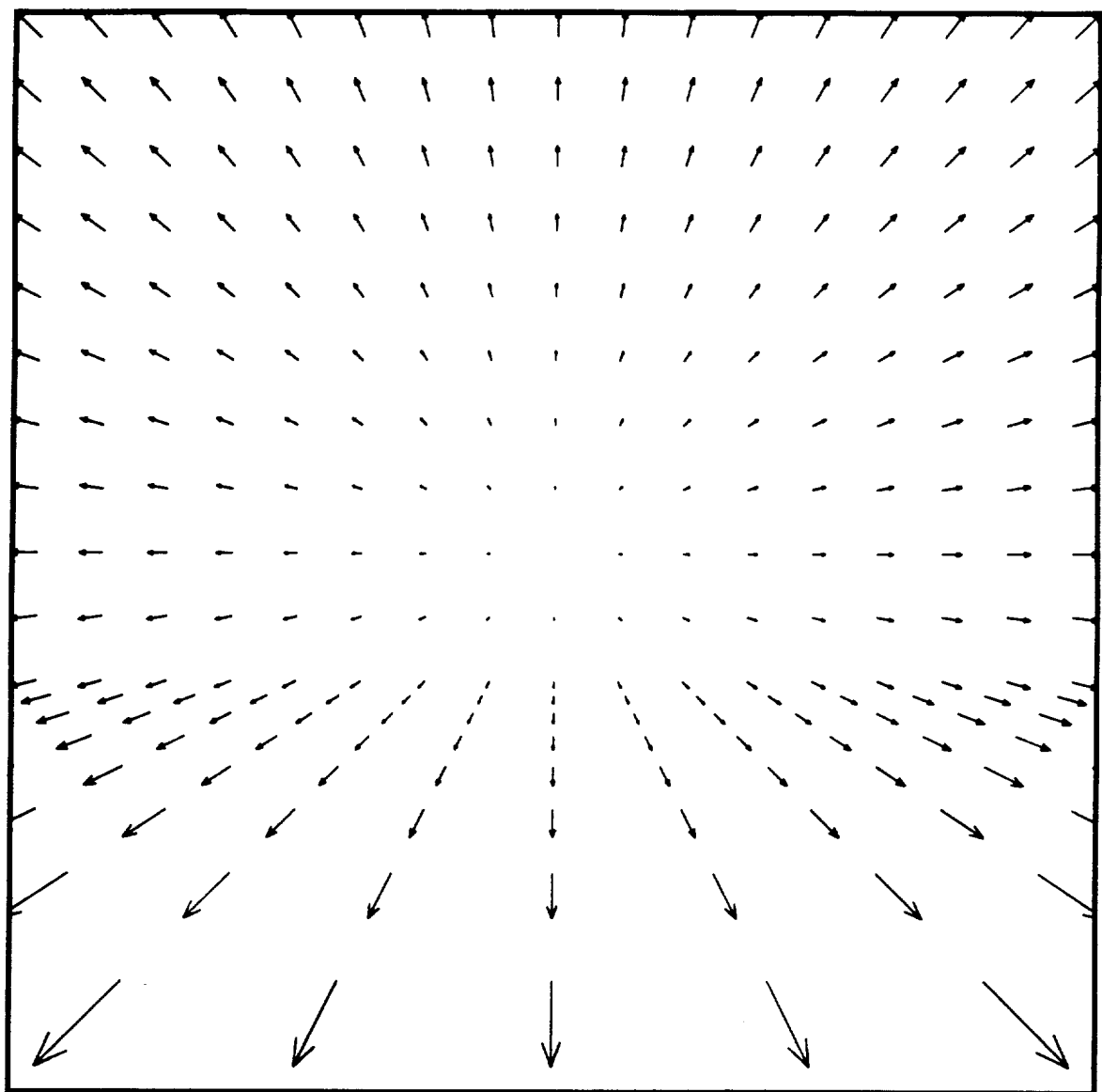Figure 4. Flow field for slanted plane when motion is parallel to image plane.

Figure 5. Flow field for motion toward a vertical surface.

OPTICAL FLOW VERSUS RETINAL FLOW AS SOURCES OF INFORMATION FOR FLIGHT GUIDANCE

James E. Cutting
Cornell University
Ithaca, New York

A recurring issue in the workshop concerned the appropriate description of visual information for flight guidance—optical flow vs. retinal flow. Most descriptions in the psychological literature are based on the optical flow. However, human eyes move and this movement complicates the issues at stake, particularly when movement of the observer in involved.

The basic question addressed here is: Can an observer, whose eyes register only retinal flow, use information in optical flow? The answer, I suggest, is that he/she cannot and does not reconstruct optical flow; instead, he/she uses retinal flow. To clarify what is meant, some definition of terms is needed.

GLOSSARY

Optical array. The projections of a three-space environment to a point within that space. All measurements to this point should be made in sterardians, solid degrees of subtended angle; typically, however, most descriptions are in degrees. The relations among these projections provide an important beginning to an understanding of information as used in visual perception. The optical array is best and most conveniently represented as a spherical projection surface, and centered on an observer's eye, which is at the nodal point of the projection.

Retinal array. The projections of a three-space onto a point and beyond to a movable, nearly hemispheric sensing device, like the retina. (1) Its movability, (2) its differential ability to register detail (acuity differences in the fovea, parafovea, and periphery), (3) its boundedness (edges at the orbit and nose), and (4) its slight deformations (due to the difference between center of rotation of the eye and its nodal point) distinguish it from the optical array. Movability is the critical factor in separation of optical and retinal flow; movability is evolutionarily designed to counter problems of acuity differences.

Flow. Global motion represented as a field of vectors, best placed on a spherical projection surface, as shown in figures 1 and 2. (These figures are Figures 11.2 and 11.3 from Cutting, 1986.) Specifically, flow is the mapping of the field of changes in position of corresponding points on objects in three-space onto a point, where that point has moved in position. Conventionally, the field of vectors is registered from two different but nearly adjacent points in three space (such as locations along the path of a moving observer). I will call these two points along the path registration points. Example of the differences between optical and retinal flow are given in figures 1 and 3 (Figure 11.4 from Cutting).

Focus of expansion. The point projected out to the horizon along the linear path one is taking. It is the (putative) point in the optical array from which all mapping vectors (flow) appear to be oriented.

Curl. The curvature of local vectors in a flow field. Some examples are given in figure 3.

Wayfinding. Determining the instantaneous direction one is traveling in.

GROUND RULES

One must be able wayfind (to determine one's heading to guide flight, particularly helicopter flight) from perceptual information. This information could be purely visual, or combined with other modalities (vestibular activity). Analyses of the wayfinding requirements during running, skiing, and landing fixed-wing aircraft (Cutting, 1986, p. 152, 277-278) suggest we need an accuracy of 1 degree of visual angle at any point in time.

Caveat: These comments do not consider overt control of an aircraft, only on the perceptual needs of the pilot to initiate control adjustments.

CAN WE DERIVE OPTICAL FLOW FROM SUCCESSIVE OPTICAL ARRAYS?

Participants at the workshop differed in their understandings about how optical flow and retinal flow information might be useful to a pilot, or any other moving observer. The issues are based on deeper conceptions of how flow in the optical array is determined. In essence, all workshop participants agree on the optical array, its generation, and its importance; there is disagreement, however, about optical flow.

In considering optical flow, should we be concerned with three degrees of freedom or six? Again, the optical array is all the projections of a three-dimensional space to a point. Since the spatial position of a point can be specified by three coordinates [x (lateral), y (vertical), and z (depth)], the optical array is concerned with only these three degrees of freedom.

However, to move the focal point of the optical array through three space, more than the change in x, y, and z may be entailed. Specifically, rotations around x (pitch), around y (yaw), and around z (roll) might have to be considered, as seen in figure 3. Here's why:

GENERAL PROBLEMS FOR CONSIDERATIONS OF OPTIC FLOW

(1) The contents of the optical array cannot be registered at a point, but only on a projection surface (hemisphere or plane) behind that point. Changes in positions of the contents of the optical array

76

need a surface on which to be represented. This means the surface of registration must be oriented to the points in the three-space from which they project.

(2) Optical flow vectors are typically extremely small within 15 degrees of the focus of expansion (see fig. 1). Threshold measures for registering motion become very important.

(3) To plot optical flow, one must map (at minimum) corresponding projected points of the environment across at least two successive optical arrays. For convenience's sake, call these registration points t1 and t2, measured at two successive time intervals. The central question is: Can these optical flow mappings be constrained on reasonable (1) mathematical, (2) optical, or (3) extraocular grounds, and/or (4) can optical flow be bypassed and wayfinding be done purely on retinal grounds?

I. MATHEMATICAL METHODS FOR CONSTRAINING OPTICAL FLOW

A purely mathematical approach can consider only projections from points in three space or from surfaces with unknown orientation and Gaussian curvature in three space, onto a spherical projection surface. In particular, the mathematics does not, a priori, allow horizons or other reference points.

1. Random choice. Pick a point anywhere in three-space and use it as the origin of the mapping system. That is, the projection of this point and only this point will necessarily map onto itself across t1 and t2 on the projection surface. It is an identity element; it will be a vector of zero length and hence no orientation. This fact is the essence of Brouwer's theorem in topology, which states that any field of mappings must have at least one point that maps onto itself. Brouwer's theorem is silent on the location of this identity element.

Result: A vector (flow) field will be generated.

Problem: Unless one has, accidentally or a priori, picked a point at (functionally) infinite distance, the flow field will have curl as shown in figure 3. Hence, this mapping will hide the location of the focus of expansion, hide the direction one is going, and give misinformation about most optic flow variables.

[Functional infinity, here, means any distance which is, say, at least three orders of magnitude larger that the distance between registration points, t1 and t2.]

2. Selection of many points and comparison of curl. Simultaneously consider many mappings using a large number of points as origins, and compare curl across mappings.

Result: Any solution without curl reveals the focus of expansion.

Problems: (1) If there are no solutions without curl, the center of the projection (the pilot's eye) is in a three-space environment with no point at (functional) infinity. In practical terms, this environment has no horizon. Nap of the earth (NOE) flight may entail such environments. (2) The procedure is either iterative (repeated stepping through comparisons of flow fields), or requires multiple coordinated registration systems (e.g. many visual systems). This is computationally expensive and

psychologically impossible. (3) If the observer pilot is on a curvilinear path, all representations of flow (optical and retinal) will have curl, and no solution can be obtained.

3. Flow decomposition. In artificial intelligence approaches to the problem, it is assumed that since contributions to flow from pitch, yaw, and roll on the one hand, and translations in x, y, and z, on the other hand are additive, the flow can be decomposed by subtractive methods. That is, any curl in the flow field occurs because of rotations and can be subtracted out.

Result and Problem. Mathematically this is problematic. It reduces to a + b = c, if a is flow due to rotations, b is flow due to translations, and c is the resultant flow. However, although knowing a and b specifies c, knowing c does not allow one to determine the values of a and b. This formulation reduces then to the iterative method above and has the same problems.

Mathematical Conclusion. Since, in mathematical terms no point in three-space can have privilege in the mapping across t1 and t2, and hence the formation of the vector field, there is no a priori, noniterative mathematical way to generate an optical flow field. The problem reduces to the fact one cannot guarantee the coordinates of the registration of flow will not have undergone pitch, yaw, or roll (rotations around x, y, and z axes), while simultaneously undergoing pure translation in x, y, and/or z.

II. OPTICAL METHODS FOR CONSTRAINING OPTICAL FLOW

Optical methods presuppose an environment. For a pilot at any substantial altitude, this environment is essentially planar and has a horizon. Any information from or about eye movements, however, is excluded in this approach.

1. Horizon. The horizon may provide one explicit method for determining the traditional representation of optical flow (e.g., figure 1), preventing unruly mappings involving "spurious" pitch (x rotation) or roll (z rotation). Because the horizon is a series of points in three space at infinite distance, any and all of these points provide anchor for flow registration across t1 and t2.

Result. Since the horizon holds a constant position in the mapping from one optical array to the next, flow vectors will not undergo curl due to pitch and roll (see the top and bottom panels of figure 3). This would facilitate location of a focus of expansion, and orientation of the optical array.

Problem. Unless texture along the horizon is used, there is no guarantee that yaw (y rotation) will not occur in flow mapping, and yaw rotations induce substantial curl, as shown in the middle panel of figure 3.

2. Horizon plus distant texture. Use the horizon to anchor the registration of flow against pitch and roll, and use any available texture near the horizon to anchor it against yaw.

Result: The result will be the "standard" optical flow pattern.

Problems: The horizon isn't usually available in some environments (inside buildings), and more particularly to the goals of the workshop the horizon is not generally available in NOE flight. It cannot be guaranteed that there will be any objects or texture at infinite distance.

3. Extrapolation from most rapid flow velocities. A look at the panels in figure 3 shows substantial curl in the vector fields near the direction of heading, but also shows large vectors (corresponding to rapid optical velocities) generally without curl beneath the individual, nearest the ground. Since these are relatively immune to curl, they could be used to determine heading.

Result: In principle this method could work and one could find one's heading perhaps within several degrees of visual angle even if the horizon were occluded.

Problems: (1) A pilot often cannot, or cannot afford to, look directly below to observe such flow, particularly in NOE flight. (2) Most rapid flow is contingent on the environment one is in and one's relation to it. In NOE flight, most rapid flow need not be beneath the aircraft and hence could be difficult to find. (3) When one "looks" at anything, one fixates on it and will use pursuit movements, creating retinal flow radically different from optical flow. To use rapid optical flow one must anchor fixation. There are three methods of doing this. Two are looking at the horizon or any other point at functional infinity. The problem here is that most rapid flow is usually at 90 degrees to this anchor, and difficult to register by eye. The third is by looking at the reference point at the edge of or on the windscreen of the craft. Such edge rates are known to be useful in open environments (e.g. over planar fields in simulations) but they would not be in NOE flight. (4) If a pilot is on a curvilinear path, the most rapid flow will have curl. Vectors will point to various locations in the distance generally in the direction of the curved path, but not along it (see Cutting, 1986, p. 209).

4. Marks on the windscreen. The best method for negating eye movements is to look at a fixed point on a windscreen and observe flow. Flow will always be in the opposite direction to one's line of movement.

Results and Problems: If one had enough marks on the windscreen one could generally pick out the focus of expansion by fixating various marks. This method is important for the practical task of piloting an aircraft, but is not available to a pedestrian or runner. Moreover, too many marks on a windscreen will impair visibility.

Optical conclusion. When one is in an environment without a guaranteed horizon, as in NOE flight, there is no optical method (other than many marks on a windscreen) that can guarantee finding the focus of expansion and anchoring the vectors in the optical flow field. Hence the standard optical flow variables are (or may be) indeterminate.

III. EXTRAOCULAR METHODS FOR CONSTRAINING OPTICAL FLOW

These methods use information from extraretinal sources, such as feedback from eye muscles and/or from the semicircular canals of the vestibular system, to anchor the registration of flow in the optical array.

1. Registering sensor rotations. Since flow due to rotations are (thought exclusively to be) due eye movements, and flow due to translation are due to observer movement, one could decompose these contributions in retinal flow by registering eye movement extent from muscle activity.

Result. In principle this would work.

Problems. In practice it does not. We are not sufficiently kinesthetically sensitive to eye movements to drive guidance within 1 degree of visual angle.

2. Preventing sensor rotations. In principle the vestibulo-ocular system is gyroscopic. The vestibular system can be used (particularly in VOR) to hold eye position, preventing rotations.

Result. Again, in principle this would work.

Problem. The normal activity of the eye is driven not by VOR but by a field holding response which serves to direct the eye to an object and hold it (or some part of it) in position on the fovea. This field holding response overrides VOR. It is extremely difficult to stare off into the distance at a fixed angle and observe flow. One is constantly captured by objects, which one pursues, then one saccades back in the opposite direction. This phenomenon is optokinetic nystagmus (OKN).

Extraocular conclusion. In principle optical flow could be determined from extraretinal sources of information. In practice, however, the information is probably too coarse, at least in human observers.

IV. A RETINAL METHOD FOR BYPASSING OPTICAL FLOW: Differential motion parallax (DMP)

A retinal method of determining flow tries to bypass problems of reconstructing optical flow. That is, rather than trying to nullify rotations of the optical array, this approach embraces the rotations in the retinal array and see what regularities fall out. In essence, the claim is that while the optical array is relevant to perception, optical flow is not; only retinal flow is relevant.

Cutting (1986, Chapters 10-13) describes a retinal invariant that serves, in most environments, to indicate the direction of forward movement, whether linear or curvilinear, with respect to gaze. Its crux is that, when one is fixating an object while moving through an environment, the retinal velocity of near objects will generally be faster and in the opposite direction to far objects. Most rapid flow is in the opposite direction from heading. Thus, regardless of one's path, so long as one is fixated on an object in mid-distance, if the most rapid flow is leftward, one heading is to the right of gaze. If most rapid flow is right, heading is left-of gaze. Thus, more simply:

$$N > -F, \qquad (1)$$

where N stands for the retinal velocities of near objects (and given positive sign) and F stands for the retinal velocities of far object. Zero retinal velocity, of course, is where one is looking.

Results of four experiments were presented in the workshop supporting DMP: three (Cutting, 1986, Chapters 12 and 13) indicate the efficacy of DMP along linear and curvilinear paths, the fourth (unpublished) indicated its efficacy in situations mimicking the bounce and sway of normal gait. Moreover, an analysis of errors indicated that DMP was used throughout, not just on correct trials. That is, given certain situations, DMP can fail due to the relative positions of objects in an environment (see Cutting, 1986, p. 197).

Retinal conclusion. Information is available in the retinal array for guidance. This information is generally trustworthy, but fails in certain environments with particular distributions of objects in it. A consideration of the failures is important to testing DMP in experimental situations.

Constraints. For DMP to operate, the minimum requirements are that objects be laid out in depth around a fixated object such that there are both nearer and farther objects near the line of sight. When there are no objects farther than the fixated object there is no problem; the motion of farther possible objects is zero and does not effect the inequality in equation 1 above. When there are no objects nearer than the fixated object, however, DMP fails completely.

What kinds of visual information DMP ignores. Differential motion parallax is measurement about pure motion. It is unconcerned with occlusions of near objects by far objects. It knows nothing about the sizes of objects, their identity, or their location in three space. It is also a measurement in the retinal array unconcerned with any ability to resolve motion. That is, it assumes retinal velocities can be measured with (roughly) equal efficacy everywhere, particularly above and below the line of gaze.

RECURSIVE RULES FOR WAYFINDING

(1) Fixate an object of potential interest in your environment.

(2) If there is no flow across the line of gaze (or vertical plane passing through the line of gaze), you are looking in the direction you are going.

(3) If there is flow across the line of gaze, follow the fixated object with pursuit eye movements,

(4) During this pursuit, register the relative motions of objects near and far around the fixated object,

(5) Assess if the information is adequate as an update of your current heading.

(6) If it is, go back to Step 1.

(7) If it is not, shift eyes in the direction opposite from the most rapid flow,

(8) Go back to Step 1.

Caveat: Despite the consideration of fixations, pursuit movements, and saccades, DMP is not an eye movement theory of wayfinding; it is a partial and corollary explanation of the efficacy of eye movements and fixations as part of visual explorations of the environment during self movement.

In particular, notice one need not implement Steps 7 and 8. Once one has the general idea of one's heading, fixations, pursuit movements, and saccades need only reinforce the perception or knowledge of heading. In other words, the pilot can implement Steps 1 through 6, never (or perhaps only occasionally) looping through the whole set.

FIVE EXAMPLES OF THE RETINAL OPTICS RELEVANT TO DIFFERENTIAL MOTION PARALLAX IN CLUTTERED ENVIRONMENTS

1. Looking at or near the focus of expansion. There is no DMP when looking at the focus of expansion. Moreover, DMP will often fail when looking near it. That is, because one wants to avoid objects in one's path, a pilot has already changed course to remove them from the path. This means there are few objects available to create the foreground motion required in DMP.

Implication: This fact, the failure of DMP near the focus of expansion, may be why pilots and car drivers spend so little time looking in the direction of their path of movement. Retinal motion information there is either (1) nil, or (2) DMP information is contradictory.

2. Looking at an object in the very far distance off the path of movement: All retinal will opposite in direction from heading. That is, leftward retinal flow indicates a gaze angle to the left of one's direction of movement; rightward flow indicates a gaze to the right. Again, technically there is no differential motion parallax, only motion perspective.

3. Looking at an object in the mid-distance: DMP reigns. Any object or texture along the line of gaze (or along a vertical plane through the line of gaze) that is half the distance (or less) to the fixated object will move faster than, and in the opposite direction to, any object or texture at infinite distance. If there are no objects nearer than half the distance, DMP will fail.

4. Looking at a close object: DMP will always fail. But looking at nearby things is generally not what one does when one is interested in where one is going.

Implication: Never look at close objects when wayfinding. This may be part of the problem with in cockpit instrumentation. Wayfinding is not about looking nearby.

5. Looking at the marks on the windscreen. Although marks on the windscreen are closer to a pilot than any external object in the environment, the fact that they do not move with respect to a pilot (sitting still) makes these marks at a distance of functional infinity. This situation becomes exactly like situation 2 above, provided the craft is not undergoing any rotations.

Meta-rules for visual guidance by DMP: Both rules have practical reasons for implementation known for along time; both are reinforced by the optics of DMP.

(1) Spend little time looking exactly in the direction one is going;
    (a) there is either no suprathreshold motion there, or
    (b) what suprathreshold motion there is may yield contradictory DMP information.

(2) Spend little time looking at objects in the near foreground;
    (a) one may have to rotate one's eye too rapidly to maintain fixation.
    (b) there can be contradictory DMP information there.

EXPLORATIONS OF DMP IN SIMULATIONS RELEVANT TO HELICOPTER FLIGHT

During the course of the workshop several of us (Tom Bennett, John Flach, Dean Owen, Lawrence Wolpert, Greg Zacharias, and I) designed an experimental situation to explore the use of DMP and optical flow. The situation uses a head-mounted display responsive to head rotations (pitch, yaw, and roll). Thus, the simulator pilot moves his/her head to obtain new vistas of the environment he/she is flying through. In this environment, one can fly a dog-leg path in NOE flight through and around a series of poles, towards a goal. Computer software is being developed to record the head movements and derive the objects the pilot is looking at. The objective, from my perspective, is to determine if pilots use DMP to guide their flight.

We assume (1) that participants can generally fly an aircraft and (2) that they can learn to use headmovements instead of eye movements. In this manner changes in the position of the center of the display should indicate the locus of the two dimensional array they find interesting. This second assumption may be prove wrong, but the participants of the workshop agreed that positive evidence of the use of DMP would also serve as positive evidence for this assumption.

REFERENCES

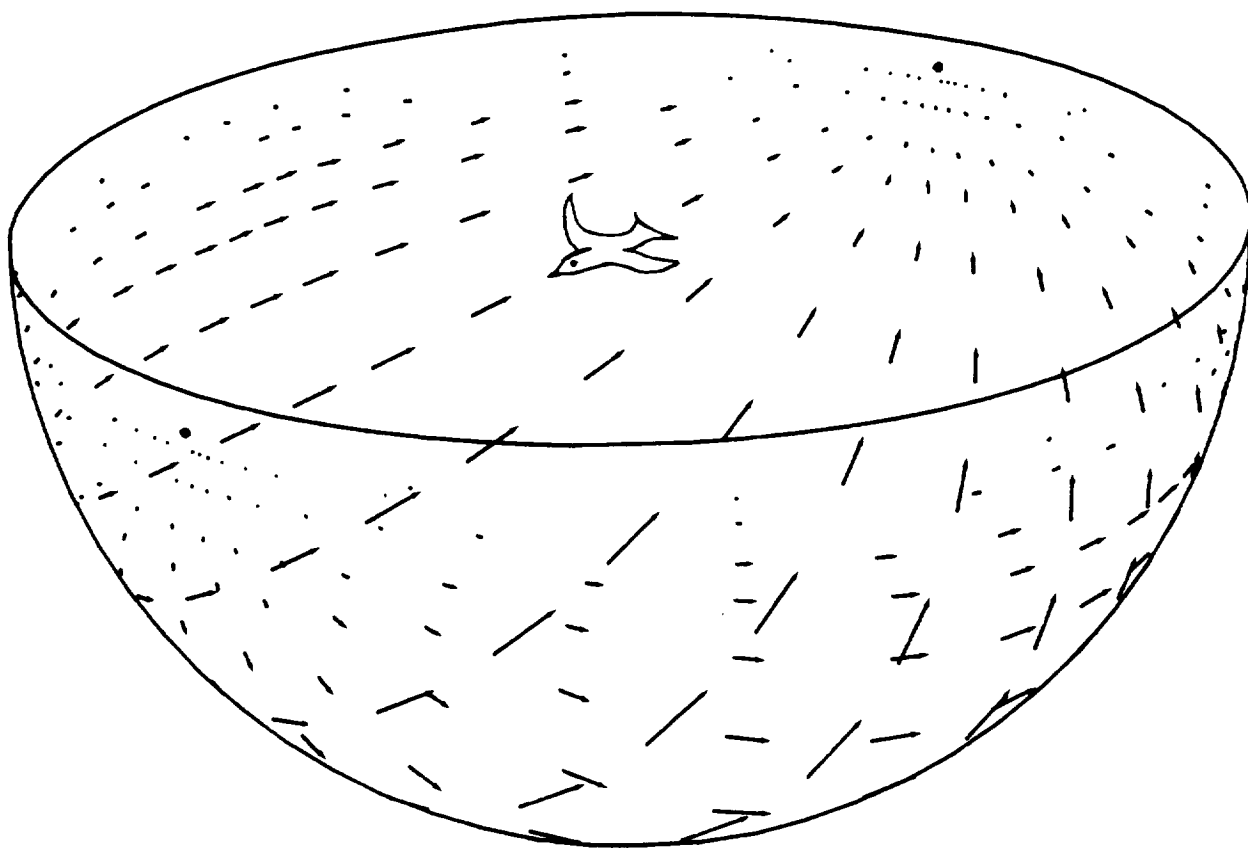Cutting, James E., Perception with an Eye for Motion. MIT Press, 1986.

Figure 1. A computer-generated plot of the projection shown in figure 11.1. Notice that the vectors near the foci of expansion and contraction are smaller than in gibson's figure.
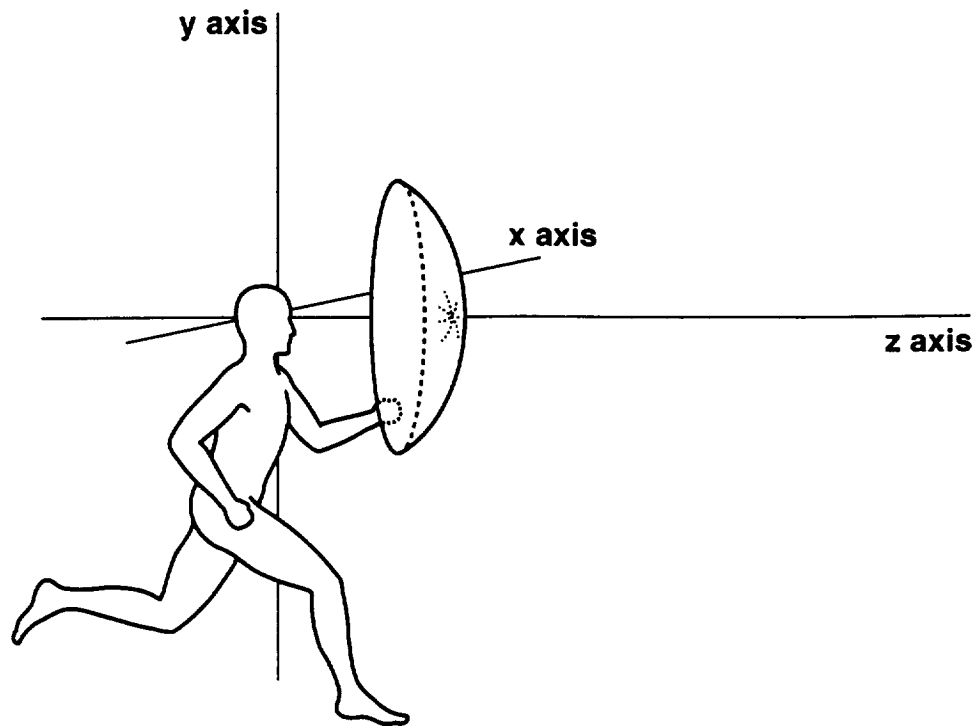
Figure 2. A representation of leonardo's window, in which the moving observer carries the projection surface through the environment. In accordance with the lower panel of figure 10.1, I call this leonardo's windshield. It is a section of the spherical projections shown in figures 11.1, 11.2, 11.4, and 11.5. Three axes of potential rotation are noted: x (which extends side to side across the observer, y (which runs vertically), and z (which extends along the linear path of movement. These axes of rotation are used in figure 11.4.
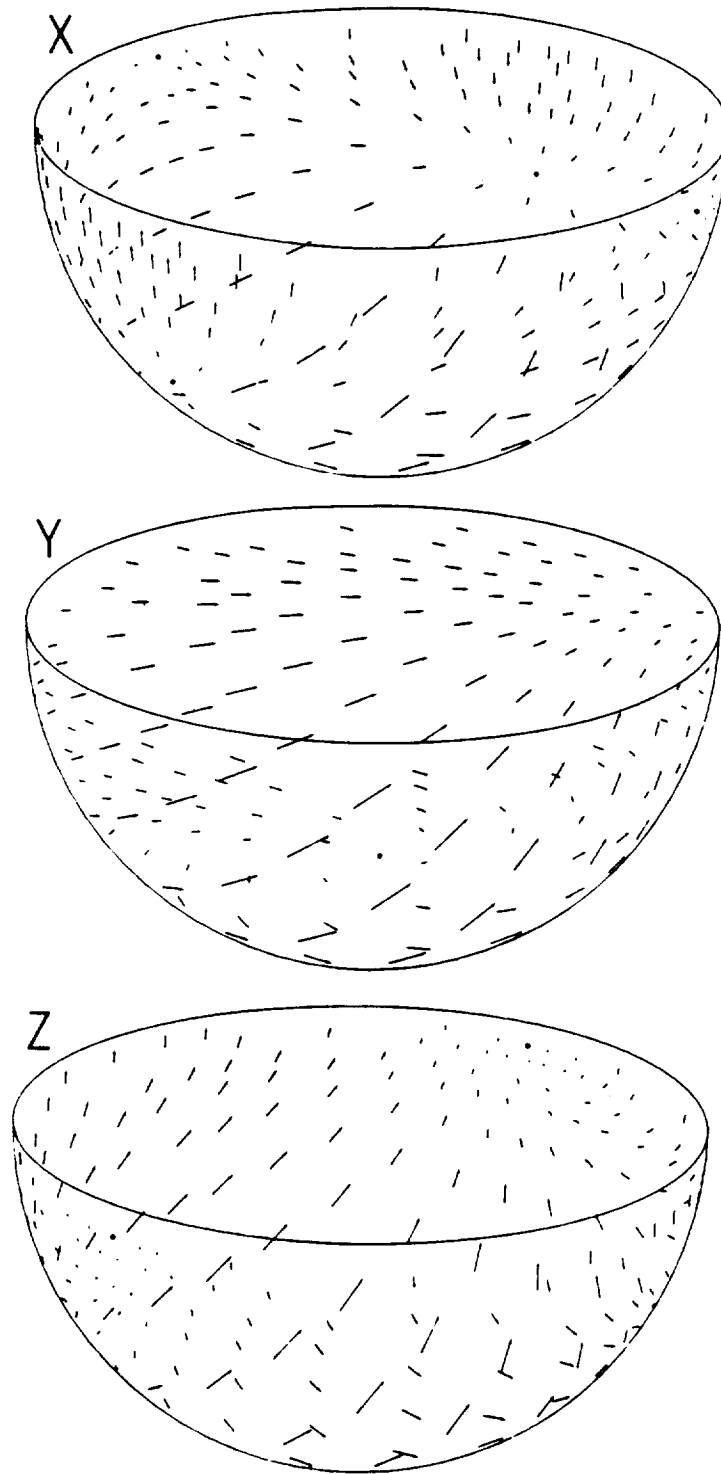
Figure 3. Alternative mappings of flow during forward linear locomotion. (Top) The spherical projection surface has been rotated around the x axis. Singularities are created where the x axis meets the horizone and beneath the observer. (Middle) Rotation around the y axis. Two new singularities are created to the left of the observer (one hidden at the edge of the drawing). (Bottom) Rotation around the z axis. The displays are valid representations of optic flow, as much so as those in figures 11.1 and 11.2.

PERCEPTION AND CONTROL OF ROTORCRAFT FLIGHT

Dean H. Owen
University of Canterbury
Christchurch, New Zealand

Of the topics receiving attention during the workshop, three overlap with my areas of expertise and interests in application to rotorcraft flight. Therefore, in this report, I will concentrate on (a) the nature of visual information, (b) what visual information is informative about, and (c) the control of visual information. The first topic generated controversy concerning what I will call the anchorage of visual perception, i.e., is it the distribution of structure in the surrounding optical array or is it the distribution of optical structure over the retinal surface? The second topic provoked debate about whether the referent of visual event perception, and in turn control, is optical motion, kinetics, or dynamics. The third issue dealt with the interface of control theory and visual perception. The relationships among these problems will constitute the organization of my report.

STIMULUS THEORY

A brief foray into stimulus theory is necessary to clarify the informative properties of stimulation. In attempting to answer what he considered to be the fundamental question for perception, i.e., "Why do things look as they do?", Koffka (1935) distinguished the proximal stimulus (the distribution of excitations to which the light rays coming from an object give rise) from the distant stimulus (the object in the geographical environment). He was concerned with functional issues, because he believed that, "as a rule,...the looks of things tell us what to do with them" (p. 76). Because he was convinced of a lack of specificity between either stimulus and the world as perceived, Koffka rejected both proximal and distal descriptions of stimulation as useful in answering functional questions in favor of a self-organizing process of field organization. Proximal and distal psychophysics persist as experimental approaches, of course, with particular concern for retinal image variables in accounts of depth, distance, and motion perception, and a variety of mediational mechanisms have replaced the field forces of Gestalt theory. Self organization continues to be an intriguing notion, but current versions consider systems the unit of analysis rather than processes, a point to which this discussion will return.

J.J. Gibson devoted much of his effort to stimulus theory and came to the conclusion, for a variety of reasons, that perceiving is anchored to the structure in the medium between the surfaces of the environment and the sensory surface. If so, he argued, the appropriate description of visual stimulation is in terms of the variables and invariants of the ambient optic array (cf. Gibson, 1958, 1961, 1966, 1979). To complete Koffka's classification system, I will call this a medial description.

Gibson considered optic array structure to be informative to an individual about the environment and the individual's relation to the environment. He proposed that visual perception is ordinarily anchored to the ambient array and that properties of optic array transformation or flow are

particularly important support for perceiving events, including motion of the self in the environment. (Note that the transformation in an array at a moving convergence point is not conceived as the difference between arrays at successive stationary convergence points, Gibson, 1966, p. 195-196.) After several attempts to answer Koffka's question, and partly as a result of defining and quantifying information in this fashion, Gibson reframed the fundamental question for perceptionists. He concluded that the primary function of perceiving is to support action, and that perceiving and acting have a reciprocal relationship. By acting, an individual produces transformations and invariants in the flow pattern that are informative about whether the actions are appropriate to achieve the intended goal. Perceiving is the active acquiring of information about which action strategy is appropriate and the relative success of behaving. Actions are initiated, modulated, and terminated in order to control the informative variables of stimulation (cf. Gibson, 1958, 1979). (An encompassing claim for this position is that phenomenal experience, nervous system activity, and performance all are anchored to the ambient optic array.) For Gibson, the problem with the highest priority is determining what properties of the ambient array are informative in each control situation.

It is important to note that optical variables are potentially informative until they are effectively informing, and that only then can they appropriately be called visual variables. As a case in point, two aircraft maintaining an invariant angle between their flight paths will collide, unless at least one of the pilots notices that the optical angle is constant over other optical transformations and initiates control adjustments to change it. The optical angle is there to be sampled; it is potentially informative about impending collision, but it is not a visual angle until a pilot samples it with a visual system, and only then is it informative.

STRUCTURE IN THE MEDIUM

In the case of self motion, the referent of perceiving is not a distant surface, but rather the relation between the moving individual and the distant surfaces. What kinds of optical support are there for detecting and controlling this relation? Three types of potentially informative medial properties can be distinguished: (a) local, (b) regional, and (c) global.

Local flow structure. Some properties of the flow pattern are available only in specific directions in the optic array. The foci of expansion and contraction are examples, and their usefulness has been controversial. Local optical density, local flow velocity, and local optical discontinuity rate are all specifiable in every location, but their regional and global gradients appear to be more informative.

Regionally distributed flow structure. Regions of the optic array are structured by (a) environmental differences and (b) visible parts of the self. The region in the direction of movement is characterized by flow expansion, the lateral regions by nearly lamilar flow, and the region opposite to the direction of movement by flow contraction. The horizon is a regional optical structure which provides an anchor for the pitch and roll dimensions of rotational self motion. The horizon also provides a referent for the optical displacement of places and objects below the horizon (the subtense or "dip" angle) and for eyeheight and change in eyeheight relative to objects extending above the horizon (the horizon ratio, cf Langewiesche, 1944; Sedgwick, 1973, 1980).

88

Since both optical density and optical flow velocity vary with the distance of surfaces from the path of motion of the eye, regions or sectors of the array structured by surfaces at different distances will reveal differences in both variables. Driving through a tunnel provides a suitably constrained case, as the regional density and regional flow velocity will vary with the distance to each of the four surrounding surfaces. Change in flow velocity, optical density, and perspectival "splay" angle (Wolpert, Owen, & Warren, 1983) all occur with change in the distance from the eye to a regional surface, so the regional character of each surface is multiply specified. Differential motion parallax arising from movement of the eye past surfaces with vertical extent is also regional (cf Cutting, 1986). By fixating a flowing optical discontinuity, the pilot is able to isolate useful flow structure in a particular region of the transforming array, and control it to achieve a goal, e.g., determining the current direction of heading or determining whether current heading is in the direction desired.

Environmental surfaces structure different regions of the optic array in different ways, but the different regional transformations and invariants can specify the same property of self motion. For example, during change in altitude, perspectival splay change is structured by ground surface texture elements, whereas change in horizon ratio and change in dip angle below the horizon are structured by surfaces with vertical extent. Since both types of surfaces are usually available during low-level flight, it is would be useful to know whether it is better to learn with redundant information, or better to learn to detect and control the various types of information separately before they are introduced in concert.

Regions of the optic array are also differentially structured by surfaces that travel in concert with the eye. These include the orbit of the eye, the side of the nose, other parts of the body which extend into the visual field, and parts of the extended ego encompassed by a moving vehicle (windscreen frame or sections of the aircraft). In the case of pure egorotation about the center of the eye, there would be no change at all in the ambient array other than that resulting from progressive occlusion of sectors of the array by the body.

Globally distributed flow structure. The defining characteristic of a global optical description is that it is independent of optical position, i.e., it is the same for every locus (Warren, 1982) and, it follows, for every region. Therefore, global array properties can be used to compare two arrays or, more commonly, to detect change in an array over time. They are especially useful and reliable because they are the same wherever the individual looks, as long as there are optical discontinuities to convey them. Some hold for both frozen and transforming arrays, and some occur only with motion. Global optical texture density, global optical flow velocity, and global optical discontinuity rate will be used as examples, since they form a linked set and have received extensive empirical attention.

Global optical texture density is defined as the number of surface texture units that can be spanned by the eyeheight of the individual (Warren, 1982). The metric is ground units per eyeheight. Since texture units are nested, a referent must be chosen for any case where more than one grain is available, e.g., fields at higher altitudes, rocks and clumps of vegetation at lower altitudes. For detection of changes in both speed and altitude, density has an optimal level, and appears to provide contextual support for other linked variables (Owen, 1989).

Global optical flow velocity is indexed by the common multiplier of path speed divided by eye-height applied to every locus in the transforming array (Gibson, Olum, & Rosenblatt, 1955; Warren, 1982). (Note that global flow velocity in eyeheights per second is equal to local flow velocity in radians per second directly below the eye.) Since global flow velocity varies with change in either speed or altitude, but not necessarily with simultaneous change in both, it is not an unequivocal specifier of either self-motion variable. Warren (1982) partitioned global optical flow acceleration into a vertical component (change in eyeheight divided by current eyeheight, i.e., fractional change in altitude) and the multiplier indexing change in flow velocity as function of change in path speed. This partitioning had two empirical consequences: (a) It was found that flow acceleration is not functionally informative about approach to the ground surface, and it in fact interferes with detection of descent (Hettinger, Owen, & Warren, 1985). (b) Fractional (as opposed to absolute) loss in altitude was found to be a functional event variable, leading to a search for functional optical variables.

Optical discontinuity rate. Optical discontinuities result from differences in surface reflectance, refraction, or emission of light. Discontinuities can be structured by elements of surface texture (e.g., rocks, trees, buildings, or dots in a schematic simulation) or by borders (e.g., edges of fields or stripes across a roadway). Discontinuity rate indexes the number of discontinuities crossing a given optical locus per unit time (Warren, 1982). Global discontinuity rate is indexed by the ratio of path speed to distance between surface discontinuities. Therefore, it depends on both egospeed and the spacing of elements or borders on the environing surfaces, but is independent of the distance of the eye from the surfaces. The role of edge rate has been studied extensively in the contexts of perceiving and controlling speed (Larish & Flach, in press) and change in speed (Awe, Johnson,& Schmitz, 1989; Denton, 1980; Owen, Wolpert, & Warren, 1984; Warren, Owen, & Hettinger, 1982; Zaff & Owen, 1987)

Fractional change. Fractional change in global flow- pattern variables have consistently proved to be the information attended to and controlled in experiments concerned with change in the direction or speed of flight. The metric is percent per second change in the variable describing the self-motion event, as well as its optical specifier. Whereas the lower-order global variables are indexed by a common multiplier on varying local properties, fractional changes are optically privileged in the global sense in that they change at the same rate at all loci. This fact may be of particular relevance to an explanation of their general salience and usefulness. Summaries of the experiments isolating the variable described above and testing their usefulness, as well as relevant references, can be found in reviews by Owen and Warren (1987) and Owen (1989).

WHAT DOES THE RETINA DO?

The relation between sensitivity to and control of the ambient flow field points toward a different conceptualization of the retina, the brain, and the rest of the nervous system than arises from mediational theories of perception and information processing theories of cognition and action in general. Most vision theorists and researchers are concerned with how the visual system recovers the nature of the visible world from retinal stimulation. If vision is instead anchored to the ambient optic array, what is the role of the retina? Gibson proposed that light is a stimulus for a rod or a cone, but not for a visual system, therefore, visual stimulation does not consist of stimuli (Gibson, 1979). Kugler and

Turvey (1987) argue that during the perceiving of an event there is a flow pattern in the nervous system. The variables and invariants of that flow are assumed to be specific to the variables and invariants of the flow pattern in stimulation.

What function does the retina have in this formulation? If the function of the entire system is specificational, then the retina must specify something about light. It would seem to have only two tasks: to specify (a) what direction the light came from and (b) what the nature of the light is. The direction from which the light comes is maintained in the curvature of the retinal surface itself. The nature of the light (frequency variation) is maintained by the selective broad-band sensitivities of the differently pigmented cells. If the retina "registers" anything, it must be these properties, but it cannot register optical flow. If the primary adaptation of the nervous system is to deal with flow fields, then it is more appropriate to consider the nervous system a medium than a processor. The retina, then, is a transducing interface between two media that support flow patterns. Is the concept of information equally at home in either flow pattern? Perhaps, but it may be more appropriate to limit informing to optical flow and consider the role of nervous system to be that of testing for reduction of uncertainty and confirming or disconfirming relative to the intended flow pattern, discrepancy from which leads to control actions modulating flow.

AFFORDANCE SPECIFICATION?

Affordances are what an individual's environment provides to support actions that result in the achievement of desirable consequences or the avoidance of undesirable consequences (Gibson, 1977, 1979). An effectivity is a set of action properties taken with reference to a set of properties of the environment which can be acted upon (Shaw & McIntyre, 1974). Gibson proposed that affordances are perceived directly on the basis of action-scaled information in the light. This concept embodies an approach to understanding what went wrong when an error is made, since it is assumed that errors are made relative to affordances. Action is scored relative to the availability of an appropriate affordance. Perception is scored correct or in correct relative to the availability of appropriate information specifying an affordance.

Affordances have consequences due to dynamics, and effectivities are also describable in terms of dynamics. A surface that affords landing upon must support the mass of the rotorcraft. To avoid colliding with the ground or objects protruding from the ground, the pilot must manage the forces under his control. These are the effectivity properties of the person-vehicle system. The argument that affordances are directly perceivable, then must entail the assumption that dynamic properties of events are perceivable. Gibson argued for a chain of specificities that links ambient-array variables with kinetics, i.e., relative motions among surfaces. Runeson extended the chain by proposing that the variables of kinematics are specific to the variables of dynamics, and conducted a series of perception experiments to support his claim that dynamics are perceivable (cf Runeson & Fryckholm, 1983, for a review). Kugler and Turvey (1987) conclude that "any flow morphology that can be defined reliably on a low energy field...is potentially a source of information about the dynamics that gave rise to it (p. 104)." Proffitt (1989a, b), in contrast, argues from the results of a series of experiments, that dynamics are not perceptually penetrable and that problems involving dynamics are solved by using unidimensional heuristics.

The experiments reported by Runeson and Proffitt have involved judgments of discrete events based on abstract knowledge. Rotorcraft flight, in contrast, involves closed- loop coupling of perception and control actions with continuous feedback from which a pilot could develop procedural knowledge. In actual flight, the pilot must deal with multidimensional dynamics, involving control, flight, and wind dynamics. If the chain of specificities is sufficiently "tight" under active control conditions, a person may learn to perceive dynamics. This learning is likely to be self organizing in that feedback is intrinsic to the extended event, so that with exploratory actions and practice, a trainee could learn without feedback from an extrinsic agent (e.g., either an instructor or a computer). If learning to fly a rotorcraft is of this type, then questions should be raised concerning how best to support self organization of the necessary skills, perhaps instead of instruction. These are problems that deserve experimental attention, and may benefit from the kind of physical theory explored by Kugler and Turvey (1987). The fact that different optical variables may be linked to the same change in dynamics might provide the needed wedge to open this issue to investigation.

CONTROL OF OPTICAL VARIABLES

The preceding discussion emphasizes the linkages among optical variables. Controlling self motion involves maintaining intended conditions of speed and direction of flight, as well as self orientation, relative to environmental surfaces. In the process, variables are linked and unlinked as speed and direction change. With knowledge of the relevance of the different kinds of information to different kinds of flight tasks, the variables and their linkages can be controlled to achieve intended goals. The same ambient array properties which were independent variables in passive judgment experiments can be recorded as dependent variables in the study of active control. This is possible for both performatory actions initiated to achieve goals or avoid problems (e.g., an undesirable collision) and exploratory actions, which may allow the individual to discover or confirm functional relationships (Flach, 1989).

"Smart" mechanisms for perception and control. It might be supposed that other flying animals have "smart" perceptual mechanisms (Runeson, 1977) for acquiring information that maps directly onto an action system specialized for controlling flight. In contrast, human flight must be mediated by a vehicle. Whereas the human's perceptual mechanisms may be sufficiently smart to pick up the relevant information, manipulation of the control surfaces is apt to be quite foreign to an animal whose effectivities and prior experiences involve adaptation to terrestrial locomotion.

Guidance of flight can be cast in terms of control of musculature or it can be described as control of the path and speed of the eyes. The latter description is equally appropriate to unmediated flight and flight mediated by a vehicle. In performing a maneuver, the pilot cycles between sampling the information available and performing control adjustments to reduce deviations from desired optical conditions, repeating the perception-action cycle until satisfactory visual conditions have been achieved. As a consequence, the information acquired by perceiving and the information controlled by acting must be the same. This linkage allows recovery of the intention of a pilot by determining the properties of the flow pattern that were invariant over segments of the flight path with which the pilot was satisfied for some duration. Control systems for vehicles have been designed primarily around engineering constraints, including those of cables, levers, and hydraulic systems. The

development of electronic and optical systems communicating between controls and control sub-systems, including power, allows for the implementation of "smart" control systems designed to provide a match between the sensitivity of the human perceptual system and the effectivities of the human-vehicle action system. Smart action systems can evolve to support flight control by other flying animals, but for human control of flight they must be developed and tested. The flight environment demands that the principles be the same. In the sections which follow, those principles will be elaborated.

Direct or "natural" control. Using the cyclic and collective, helicopter pilots currently make an average of 50 control adjustments per minute during an approach to hover above a place on the ground. Pilots are instructed to keep "visual streaming" constant at the rate of a brisk walk during an approach to hover. Traditional controls usually operate mechanical linkages or hydraulically actuated systems to change an effector (control surface or power source). Recent fly-by-wire and fly-by-light technology allows interfacing a computer between the control and the effector. The computer can take inputs from the control and sensors (e.g., radar altimeter, forward-looking radar, a signal transmitted from the ground or a ship) and make adjustments in speed and direction that match the differences in event properties perceived or intended by the pilot. For approach to the ground or to surfaces with vertical extent, a fractional rate controller can reduce speed in the same proportion as distance to the surface is decreased. The pilot selects a fractional rate which matches the task demands, e.g., a high rate when time is critical, a low rate when accuracy is important. A second mode of control is appropriate for path angle. Whereas magnitude controllers vary the numerator or denominator of the ratio of vertical speed to ground speed, a path-slope controller varies the ratio directly. Since path slope equals the "dip" angle of the point of optical expansion below the horizon, the path-slope controller gives the pilot control over what he intends to achieve visually. Similar ratio modes could be developed for rotational control.

A control system designed around perception-action compatibility should reduce flight-control demands, freeing the pilot's attention for other workload. Maneuvers under difficult conditions should be simplified. Given that control is scaled in units of distance to the ground, fractional-rate control is particularly appropriate to low-level contour and terrain following. A design criterion for some new aircraft is that "trainability" be taken into account during development of the aircraft itself. Ratio controllers are relevant to this criterion, since training should be considerably simplified with a high compatibility system having independent modes of control, as compared to the current system involving complicated and sometimes arbitrary relationships between control adjustments and visual stimulation as well as interdependent relationships between the controls themselves. The proposed modes of control should also greatly simplify training and increase safety at low altitudes in cluttered environments and under difficult conditions, e.g., high work load or stress. Although experienced helicopter pilots have shown no sign of negative transfer, having a computer in the control loop means that traditional modes of control could be selected by a pilot who was trained with those modes.

It is important to emphasize at this point that the entire system should be the unit of analysis, rather than studying perception and control separately. A particular mode of control may be best given a particular kind of optical information, so that the adequacy of a control mode may vary with task and environmental conditions. The relevant interactions cannot be investigated without simultaneously varying kinds and distributions of surface texture, information acquisition strategies, and

modes of control. These variables may also affect transfer of training and transfer of research findings from simulation to actual flight by interacting with types of simulation, i.e., a window on the head (head mounted display), a window on the vehicle, or a window on the world (dome display representing a sector of the ambient array).

REFERENCES

Awe, C. A., Johnson, W. W., and Schmitz, F. (1989). Inflexibility in selecting the optical basis for perceiving speed. Paper presented at the 33rd Annual Meeting of the Human Factors Society.

Cutting, J. E. (1986). Perception with an eye for motion. Cambridge, MA: MIT Press.

Denton, G. G. (1980). The influence of visual pattern on perceived speed. Perception, 9, 393-402.

Flach, J. M. (1989, October). Exploratory behavior in the control of egomotion. In L. J. Hettinger (Chair), Visually guided control of self motion. Symposium conducted at the 33rd Annual Meeting of the Human Factors Society.

Gibson, J. J. (1958). Visually controlled locomotion and visual orientation in animals. British Journal of Psychology, 49, 182-194.

Gibson, J. J. (1961). Ecological optics. Vision Research, 1, 253-262.

Gibson, J. J. (1966). The senses considered as perceptual systems. Boston: Houghton Mifflin.

Gibson, J. J. (1977). The theory of affordances. In R. Shaw and J. Bransford (Eds.), Perceiving, acting, and knowing: Toward an ecological psychology (pp. 67-82). Hillsdale, NJ: Erlbaum.

Gibson, J. J. (1979). The ecological approach to visual perception. Boston: Houghton Mifflin.

Gibson, J. J., Olum, P., and Rosenblatt, F. (1955). Parallax and perspective during aircraft landings. American Journal of Psychology, 68, 372-385.

Gilden, D. L., and Proffitt, D. R. (1989). Understanding collision dynamics. Journal of Experimental Psychology: Human Perception and Performance, 15, 372-383.

Hettinger, L. J., Owen, D. H., and Warren, R. (1985). Global optical flow pattern information for loss in altitude. In D. H. Owen (Ed.), Optical and event-duration variables affecting self-motion perception (AFHRL-TP-85-23, AD-A161-836, pp. 155-176), Williams AFB, AZ: Operations Training Division, Air Force Human Resources Laboratory.

Koffka, K. (1935). Principles of Gestalt psychology. New York: Harcourt, Brace.

Kugler, P. N., and Turvey, M. T. (1987). Information, natural law, and the self-assembly of rhythmic movement. Hillsdale, NJ: Erlbaum.

Langewiesche, W. (1944). Stick and rudder. New York: McGraw Hill.

Larish, J. F., and Flach, J. M. (in press). Sources of optical information useful for the perception of velocity of rectilinear self-motion. Journal of Experimental Psychology: Human Perception and Performance.

Owen, D. H. (1989). Perception and control of changes in self motion: A functional approach to the study of information and skill. In A. Wertheim and R. Warren (Eds.), Perception and control of self motion and orientation (pp. 265-303). Hillsdale, NJ: Erlbaum.

Owen, D. H., Wolpert, L., and Warren, R. (1984). Effects of optical flow acceleration, edge acceleration, and viewing time on the perception of egospeed acceleration,. In D. H. Owen (Ed.), Optical flow and texture variables useful in detecting decelerating and accelerating self-motion (AFHRL-TP-84-4, AD-A148 718). Williams, AFB, AZ: Operations Training Division, Air Force Human Resources Laboratory.

Owen, D. H., and Warren, R. (1987). Perception and control of self motion: Implications for visual simulation of vehicular control. In L. S. Mark, J. S. Warm, and R. L. Huston (Eds.), Ergonomics and human factors: Recent research and advances. New York: Springer-Verlag.

Proffitt, D. R., and Gilden, D. L. (1989). Understanding natural dynamics. Journal of Experimental Psychology: Human Perception and Performance, 15, 384-393.

Runeson, S. (1977). On the possibility of "smart" perceptual mechanisms. Scandinavian Journal of Psychology, 18, 172-179.

Runeson, S., and Frykolm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. Journal of Experimental Psychology: General, 112, 585-615.

Sedgwick, H. A. (1973). The visible horizon: A potential source of visual information for the perception of size and distance. (Doctoral dissertation, Cornell University) Dissertation Abstracts International, 34, 1301B-1302B. (University Microforms No. 73-22,530).

Sedgwick, H. A. (1980). The geometry of spatial layout in pictorial representation. In M. A. Hagen (Ed.), The perception of pictures (Vol. 1, pp. 33-90). New York: Academic Press.

Shaw, R., and McIntyre, M. (1974). Algoristic foundations to cognitive psychology. In W. B. Weimer and D. S. Palermo (Eds.), Cognition and the symbolic processes (pp. 305-365). New York: Erlbaum/Wiley.

Warren, R. (1982). Optical transformations during movement: Review of the optical concomitants of egomotion. (Final Technical report for Grant No. AFOSR-81-0108). Columbus, OH: The Ohio State University, Department of Psychology, Aviation Psychology Laboratory.

Warren, R., Owen, D. H., and Hettinger, L. J. (1982). Separation of the contributions of optical flow rate and edge rate on the perception of egospeed acceleration. In D. H. Owen (Ed.), Optical flow and texture variables useful in simulating self motion (I) (Interim Tech. Rep. for Grant No. AFOSR-81-0078, pp. D-1 to D-32). Columbus, OH: The Ohio State University, Department of Psychology, Aviation Psychology Laboratory.

Wolpert, L., Owen, D. H., and Warren, R. (1983, June). The isolation of optical information and its metrics for the detection of descent. In D. H. Owen (Ed.), Optical flow and texture variables useful in simulating self motion (II) (Final Technical Report for Grant No. AFOSR-81-0078). Columbus, OH: The Ohio State University, Department of Psychology, Aviation Psychology Laboratory.

Zaff, B. S., and Owen, D. H. (1987). Perceiving and controlling changes in the speed of self motion. In D. H. Owen (Ed.) Optical and event-duration variables affecting the perception and control of self motion (Final Technical Report for AFHRL Contract No. F33615-83-K-0038). Columbus, OH: The Ohio State University, Department of Psychology, Aviation Psychology Laboratory.

AN INFORMAL ANALYSIS OF FLIGHT CONTROL TASKS

George J. Andersen
University of California
Riverside, California

An important theme of this workshop has been to bring together experts from several different domains to discuss issues important in rotorcraft flight control. The experts have come from several different domains including psychophysics, control theory, human factors, and engineering. One goal of this workshop was to use interactions among the experts in these different domains as a way to understand problems in flight control. While the majority of these interactions, in my opinion, have been successful I believe that the workshop as a whole focussed to much on the specifics and may have missed the big picture of how these different areas are relevant to flight control. In this paper I will suggest a perceptual description of what I believe to be the major issues in flight control. Although this opinion will be from the viewpoint of a psychophysicist I hope that it captures the importance of some the issues from other research domains represented by others who attended the workshop.

When one considers the task of a pilot controlling a helicopter in flight, we can decompose the task in several subtasks. These subtasks include (1) the control of altitude, (2) the control of speed, (3) the control of heading, (4) the control of orientation, (5) the control of flight over obstacles, and (6) the control of flight to specified positions in the world. The first four subtasks can be considered to be primary control tasks as they are not dependent on any other subtask. However, the latter two subtasks can be considered hierarchical tasks as they are dependent on other subtasks. For example, we can decompose the task of flight control over obstacles as a task requiring the control of speed, altitude, and heading. Thus, incorrect control of altitude should result in poor control of flight over an obstacle.

The following sections will discuss each of these task separately. Within this context the importance of possible perceptual information will be discussed.

1. The control of altitude.

Of all the tasks outlined above the control of altitude is one which has received the greatest empirical investigation as a flight control task. Warren has proposed that splay rate (the change in the angle formed by meridian lines converging at the horizon) is a useful source of information whereas Owen has proposed that edge rate (the number of texture elements that pass a specific location in the visual field) is a useful source of information. It is important to note that the effectiveness of these sources of information are dependent on specific constraints present in the world. Specifically, splay rate is only useful if the meridian lines are parallel in the world. Edge rate requires that texture elements be stochastically distributed evenly in the world. While the effectiveness of these sources of information have been investigated in several studies, it is important to realize that they also require that the world be flat and rigid. It is likely that for flight control over varying terrain

other sources of information, such as the slant of surfaces, the speed of translation, or the absolute distances between the aircraft and points in the world be recovered.

2. The control of speed.

The control of speed is another task that has been studied extensively. The sources of information that have been investigated for this task are edge rate and global optic flow rate. Again, it is important to note that the use of these sources will only be effective if the simulated world is flat and rigid. For flight control over varying terrain it may be necessary to recover information for determining altitude, slant of surfaces, and absolute distances to points in the world.

3. The control of heading.

There are two sources of information that have been proposed to be useful for the control of heading--the focus of expansion (or the point of maximum divergence) and differential motion parallax. Gibson was the first to suggest the usefulness of the focus of expansion and this has resulted in many computational analyses (Lee, Perrone, Koenderink) which use this source to extract out other characteristics of the environment (e.g. relative depth and time to contact). Differential motion parallax (the different rates of velocity of points moving above and below a point of fixation) was proposed by Cutting. A considerable body of research has been conducted to determine what information is used by human observers. Johnston, White and Cummings found that subjects could not determine the focus of expansion for displays simulating motion towards a frontal parallel surface. Warren, found that subjects were accurate in determining heading for displays simulating motion to a ground plane where the direction of looking was decoupled from the direction of motion. Regan and Beverely, found that subjects could not determine the focus of expansion for displays simulating motion towards a frontal parallel surface when a simulated eye fixation was included in the transformation.

However, Reiger and Toet found that subjects could determine direction of heading for displays simulating motion towards frontal parallel surfaces. In their research subjects were quite accurate when the display simulated motion towards two overlapping transparent frontal parallel surfaces that were separated in depth. However subjects were inaccurate when the display contained only a single frontal parallel surface. Finally, work by cutting found that subjects could determine the direction of heading for displays simulating motion through an array of poles that were positioned at varying simulated depths from the observer. Although an initial inspection of the literature would suggest that the results from several studies are contradictory, a closer inspection of these studies suggests an interesting pattern---those studies that failed to find good accuracy involved displays that did not have variations in depth whereas those studies that found good accuracy did involve variations in depth. This suggests that differential motion parallax, which is only effective if the display contains variations in depth, may be the source of information used by human observers.

It is important to note that differential motion parallax and the focus of expansion can only be extracted for rigid worlds. Constraints such as altitude, speed, or absolute distance are not required to use either source of information.

4. Control of orientation.

The control of orientation has traditionally been decomposed into the control of roll (rotation about the line of sight), pitch (rotation about the horizontal axis), and yaw (rotation about the vertical axis). In order to accurately control roll and pitch there are two sources of information that could be used---a change in the direction of the gravitoinertial vector, and a change in the horizon. The direction of the gravitoinertial vector can only be estimated by nonvisual sensory systems such as the vestibular and kinesthetic systems. However, changes in the horizon can be determined by the visual system. In order to control yaw information from the vestibular and visual systems can be used if the rotation involves an acceleration of deceleration.

Another issue of importance for determining changes in orientation is the need to use a frame of reference. A frame of reference (such as the frame in the rod and frame effect) can be viewed as providing information regarding a false horizon.

An alternative way to consider the control of orientation is to describe orientation change as a change in the position of the viewer with respect to the environment. This definition allows us to consider orientation as an issue regarding locomotion in the world as opposed to rotation in the world. I believe this definition is useful as it incorporates navigational issues (such as where am I located on this map) which are extremely important for nap of the earth and low level flight. Inconsistencies between where you think you are when you look out of a cockpit and where you think you are when viewing a map may lead to disorientation.

5. Control of flight over obstacles.

The control of flight over obstacles is an issue that has not received much attention. This is probably a result of the fact that accurate control of flight over obstacles requires the integration of several sources of information. At a minimum it requires information regarding altitude, speed of motion, and heading. In addition, it may require other information such as time to contact, absolute distance to surfaces in the environment, the slant and elevation of the obstacle to be flown over, and the location of the horizon. In many respects these sources of information may be interrelated. For example, a misperception of where the horizon is located may result in a misperception of slant. This could result in a misperception of elevation of the obstacle to be flown over which may have drastic effects on the ability of a pilot to successfully fly a nap of the earth mission.

6. Control of flight to targets.

The control of flight to targets is another type of flight control task that has not received much attention in the literature. There are two versions of this type of flight control that should be considered. One version involves the control of flight to a target that is visible from the outside scene. The second version involves the control of flight to a specific target when the pilot can not see the target in her field of view but has a map which indicates that location of the target in the world.

For the control of flight to a target visible from the outside scene there are several sources of information that the pilot must use. The pilot must determine the difference between the current heading of the helicopter and the desired heading to the target. In addition the pilot must determine

101

the current speed of travel in order to produce an appropriate control adjustment for approach to the target.

For the control of flight to a target not visible in the scene the pilot must navigate such that her current position changes in accordance with a desired location in the world. This task not only requires that the pilot's perceived location in space (from information in the visual world) match the perceived location of the pilot's position on a map but also requires that the pilot correctly determine the relative position of landmarks in the visual scene. Incorrectly perceiving the layout of these landmarks most probably result in poor flight control through these landmarks.

One interesting issue regarding this flight control to a target is whether the pilot must recover the spatial layout of the world in order to successfully perform this task. It may be that all that is necessary to correctly perform this type of task is to recover the spatial layout of the landmarks rather than the spatial layout of the world with the relative position of the landmarks nested in the layout of the world.

SENSITIVITY TO EDGE AND FLOW RATE IN THE CONTROL OF SPEED AND ALTITUDE

Lawrence Wolpert, Ph.D.
Logicon Technical Services, Inc.
Dayton, Ohio

A number of studies have examined the potential efficacy of global optical flow rate and edge rate for specifying changes in self-motion. These have ranged from passive judgments of simulated accelerating self-motion to the active control of altitude in the presence of changes in flow and edge rates. This report will summarize a number of these studies and attempt to reconcile their respective findings.

Edge rate, defined as the number of texture edges traversed per unit time, was studied by Denton (cf. 1980), first in a simulator and then on an actual roadway. Using an automobile simulator, he found that he was able to manipulate subjects' control of forward speed by spacing texture edges on the roadway at decreasing intervals. While the task was to maintain a constant forward speed, the resultant increase in edge rate caused the subjects to reduce their speed inappropriately.

In contrast to edge rate, which is dependent on one's forward velocity and the spacing of texture edges on the ground, global optical flow rate depends on one's forward velocity and instantaneous altitude, and is independent of the texture density over which one is travelling. Warren, Owen, and Hettinger (1982) and Owen, Wolpert, and Warren (1983) examined the effects of gains in edge and flow rate by manipulating the spacing of edges and the velocity with which observers traversed those edges during simulated level flight. Subjects were instructed to make judgments of acceleration and were found to be differentially sensitive to these two sources of information. While some observers were sensitive to the increase in edge rate, others were not affected by edge spacing at all, and were almost entirely sensitive to increases in optical flow.

Awe, Johnson and Schmitz (1989) questioned whether people could use flow rate information to control speed in an active control paradigm. Their subjects were instructed to attend to flow rate or edge rate information, or both, and to maintain a constant forward velocity. Even though feedback was provided, subjects continued to use edge rate information as the basis for controlling their forward speed in all conditions, including the flow rate one. This was interpreted as evidence of inflexibility in selectively attending to information for self speed.

In another "active" test of the effect of flow rate and edge rate, Wolpert, Reardon, and Warren (1989) required subjects to maintain a constant altitude in the presence of changing flow and edge rates. Increases and decreases in flow rate were effected by the use of a simulated accelerating tailwind or headwind, respectively, while the corresponding changes in edge rate were obtained by manipulating the spacing of edges over which the trials were flown. It should be noted that had the subjects not touched the control stick during the trial, altitude would have remained perfectly level with the exception of a minor, zero-mean disturbance due to the windgust. It was hypothesized that increasing optical flow during level flight would lead the flow-sensitive individuals to perceive a loss

in altitude and would result in a compensatory action, i.e., an attempt to increase altitude. Conversely, on encountering decreasing optical flow, flow-sensitive individuals would reduce their altitude in an attempt to hold optical flow constant. Changes in edge rate should not have had any effect in altitude since edge rate is defined independent of altitude, and, only had subjects confused edge rate with flow rate, would we have expected results similar to those hypothesized for flow rate change.

Twenty naive subjects viewed the simulated scenes representing flight at an initial altitude of 64 feet over flat, rectangular fields. The texture pattern was made up of a black grid laid over a green world and displayed on a 90-deg wide projection screen. A pseudorandom windgust consisting of a sum of five sine waves with a mean rms error of 0 was used as a forcing function in the vertical dimension. The forcing function repeated itself four times over the course of the trial and remained in effect for its 25-s duration. Proportional change in flow rate (Rx' = 0.95, 1.00, and 1.05), was partially crossed with three levels of the second factor, proportional change in edge rate (RE' = 0.95, 1.00, and 1.05). The cells, Rx' = 0.95, RE' =1.05 and Rx' = 1.05, RE' =0.95 were omitted to yield seven events.

A number of dependent measures were recorded and analyzed. These included mean altitude, root mean square error in altitude, absolute (unsigned) error, and standard deviation in altitude over the entire trial. In addition, each trial was divided into four equal segments of 256 frames each, and the above measures calculated per bin.

Proportional change in flow rate (Rx') was significant ($p < 0.0005$) and accounted for 3.4% of the variance in altitude. Mean altitude rose from 65.3 ft at the Rx'=0.95 level to 74.4 ft at the Rx'=1.05 level. Similarly, RMS error, absolute (unsigned) error, and standard deviation in altitude grew significantly with increased proportional changes in flow rate. In contrast, proportional change in edge rate, while significant in terms of mean altitude ($p < 0.001$), accounted for only 0.8% of the variance in that measure. Mean error in altitude increased from 69.0 ft to 70.1 ft for RE'=0.95 and RE'=1.05, respectively.

When the time histories were divided into four equal temporal quarters, this variable had a significant main effect ($p < 0.0001$, R2=4.5%) as indexed by mean altitude, which increased from 66.0 ft in the first segment to 72.8 ft in the fourth. This variable also interacted with proportional change in flow rate ($p < 0.0001$, R2=2.3%). A proportional gain in flow rate, i.e., Rx' = 1.05, led to an increase in altitude from 66.2 ft in the first segment to 83.6 ft by the fourth segment. A proportional loss in flow rate, i.e., Rx'= 0.95, resulted in a decrease in altitude from 65.5 ft at the beginning of the trial to 64.4 ft at the end, while a constant flow rate (Rx' = 0.0) produced intermediate performance.

It should be reiterated that all the above results are "illusory" in the sense that, had the subject not touched the control stick at all during the event, altitude would have remained perfectly level except for the zero-mean windgust.

The fact that proportional change in optical flow had a much stronger effect than proportional change in edge rate on altitude control, (i.e., more than 4 times as much variance was accounted for), is interesting for a number of reasons. Firstly, while earlier passive studies (e.g., Owen, Wolpert, & Warren, 1983) had shown edge-rate gain to have a much stronger effect than flow- rate gain on

"acceleration" reports, the latter was much more effective in "driving" altitude in the current experiment. Subjects were more susceptible to an illusory change in altitude when flow rate was increased or decreased, than when edge rate was proportionally modified. This was more noticeable when flow rate increased rather than decreased; their altitude was perceived as decreasing and the resultant compensatory control action led to an increase in altitude.

Secondly, slightly more observers in the passive "acceleration" study (Owen, et al., 1983) proved to be "edge-rate" sensitive than "flow-rate" sensitive. In the current study, 17 of the 20 subjects showed a heightened sensitivity to gains in flow rate rather than edge rate, and 12 of the 20 to losses in flow versus edge rates. This was probably due to the nature of the task. While the former study simulated level flight and the observer was required to detect "acceleration", the present study required the subject to maintain a constant altitude and no control over forward velocity was enabled. Since edge rate typically covaries with flow rate during level self motion, equal distributions of observer sensitivity are expected when the task demands a forward-velocity-related report or action. During altitude change, however, edge rate over regular texture remains constant while flow rate usually varies, so an increased sensitivity to proportional changes in this optical variable would be anticipated. This effect was obtained in the current study, albeit only for increases in flow rate. In fact, there was a tendency over the entire experiment to gain altitude during the trial, and in only a few trials was altitude "driven" downward. This bias could be considered as an attempt to maintain a "margin of safety" but needs to be further examined, i.e., by beginning the trial at a higher initial altitude.

How can the different sensitivities, i.e., to edge rate in the Awe et al. (1989) study, and to flow rate in the Wolpert et al. (1989) study be reconciled? Why were subjects in the former unable to hold flow rate constant even when instructed to do so, while in the latter study, flow rate had a much greater effect than edge rate in "driving" altitude? A speculative answer, perhaps, lies in the relationship between the independent variables and the dependant variables in the respective experiments. In the Awe et al study, altitude was held constant while subject were asked to control either optical flow (x'/z) or edge rate (x'/xg). Since altitude (z) was fixed and edge spacing (xg) was controlled by the experimenter, any control the subject exercised was necessarily on speed (x'). In the Wolpert et al study, on the other hand, forward velocity was under the experimenter's control, while the only degree of freedom available to the subject was in the altitude dimension. Since the altitude component is present in the optical flow notation but not in the edge rate notation, it is plausible that optical flow would be the dominant variable in this form of self-motion study. During level self-motion, both flow rate and edge rate covary, differing by a scale factor. In the absence of the altitude component, edge rate, comprised of edge spacing and the change of edge spacing, would dominate.

While the above explanation is admittedly speculative, a more rigorous test of this hypothesis would allow the subject control over both altitude and forward velocity and require the maintenance of a constant altitude and/or a constant flow or edge rate. By recording performance in both the altitude and the forward velocity domains, a better understanding of the individuals' sensitivities would be obtained.

REFERENCES

Awe, C., Johnson, W. W., & Schmitz (1989). Inflexibility in selecting the optical basis for perceiving speed. Paper presented at the 33rd Annual Meeting of the Human Factors Society.

Denton, G. G. (1980). The influence of visual pattern on perceived speed. Perception, 9, 393-402.

Owen, D. H., Wolpert, L., & Warren, R. (1983, November). Effects of optical flow acceleration, edge acceleration, and viewing time on the perception of egospeed acceleration. In D. H. Owen (Ed.), Optical flow and texture variables useful in detecting decelerating and accelerating self motion (Interim Technical Report for AFHRL Contract No. F33615-83-K-0038). Columbus, OH: The Ohio State University, Department of Psychology, Aviation Psychology Laboratory.

Warren, R., Owen, D. H., & Hettinger, L. J. (1982). Separation of the contributions of optical flow rate and edge rate on the perception of egospeed acceleration. In D. H. Owen (Ed.), Optical flow and texture variables useful in simulating self motion (I) (Interim Tech. Rep. for Grant No. AFOSR-81-0078, pp. D-1 to D-32). Columbus, OH: The Ohio State University, Department of Psychology, Aviation Psychology Laboratory.

Wolpert, L., Reardon, K., & Warren, R. (1989). The effect of changes in edge and flow rates on altitude control. Proceedings of the Fifth International Symposium on Aviation Psychology, pp.749-754. Columbus, OH: The Ohio State University, Dept. of Aviation

MODELING THE PILOT IN VISUALLY CONTROLLED FLIGHT

Walter W. Johnson
NASA Ames Research Center
Moffett Field, California
and
Anil V. Phatak
Analytic Mechanics Associates
Sunnyvale, California

INTRODUCTION

Numerous experiments have been performed to determine the transfer function for human operators in simple instrument-based feedback control tasks. For example, the simplest model for a human operator is a gain with a time delay, (which usually ranges between 0.15 and 0.4 seconds). However, there have been no comprehensive studies evaluating human control strategies in visually controlled flight (i.e. flight using a visual scene and not instruments.) This paper describes the results of preliminary studies on this topic.

Human visually guided flight control is important both in low level flight, where it predominates, and in higher altitude flights, where instrument failure is always a potential danger. Researchers have applied two general approaches to this problem, one founded in high order perceptual psychophysics, and the other in control systems engineering. These are described below.

PSYCHOPHYSICAL APPROACH

The psychophysical approach examines what complex optical or perspective relationships people use in self-movement perception, and their sensitivity to such variables. The visual scene is a segment of an optic array, which, in turn, is the two-dimensional perspective mapping of the three-dimensional world onto an observation point. This visual scene may be characterized as an array of varying intensity or brightness levels rich in relationships which inform the observer about his orientation and movement (e.g., see [1] for a discussion of some of the cues that are available in a visual flight task). Humans not only can perceptually identify and extract basic optical features such as points and edges, but, they also can directly extract and regulate significant higher-order features such as optical texture size, optical shapes, and spatio-temporal patterns. According to Gibson [2], the optic array contains important features or cues that are directly regulated or controlled during flight. Furthermore, these cues may be related to aircraft state variables in only complex and indirect ways. However, little is known about how humans use these cues for vehicular control.

Unfortunately, it is unclear how well this approach accounts for manually-controlled flight, since perceptual psychologists have typically left the study of active control to the engineering

community. Furthermore, the psychophysical approach is in direct contrast with the assumption, embodied in many engineering approaches, that pilots operate upon a recovered representation of aircraft states and environmental disturbances, not upon the raw perception. Instead, engineering approaches assume that humans rely on optical variables to retrieve estimates of these state variables, which are, in turn, used for control. Thus the engineering approach has produced control laws which do not reflect control activity that is guided directly by optical variables or patterns.

ENGINEERING APPROACH

An examination of engineering approaches for analyzing visually- controlled flight reveals two significant threads. One is the use of classical control methodology to describe simply the input/output behavior of control systems. This thread relies minimally on psychological assumptions and is represented best by the classical input/output quasi-linear describing function representation or model [3]. The other thread is the use of substantial theoretical assumptions about human behavior, in combination with modern control theoretic techniques, to construct models. This thread is represented best by the optimal control model, which is based on a linear, quadratic, gaussian (LQG) optimal control formulation [4]. The describing function approach treats human control as a "black box" problem, and concentrates on measuring and representing input/output relationships. In contrast, the optimal control model formulation encompasses a psychological model which decomposes human control strategy into two cascaded processes operating on the raw input variables.

The optimal control model assumes that humans first process raw perceptual input through a Kalman filter which yields estimates of vehicle and disturbance states. This model also assumes that humans have internal models of the vehicle dynamics and the disturbance inputs that can be represented mathematically in a common, earth-fixed inertial frame of reference. The model also assumes that humans operate upon these estimates using an optimal linear quadratic controller. Application of this model to visual control tasks uses image features or optical variables as the input variables, but then gives these to a Kalman filter for estimating the vehicle and disturbance states. It is these estimated states, and not the optical variables, which are then controlled. This is assumed to be accomplished with a linear full state feedback controller designed to minimize a quadratic cost function.

Thus, modern control theory and the psychophysical approach represent directly competing models of the information humans might use to control flight. The optimal control model presumes that a non-optical frame of reference is used by humans. It poses the control problem as being, in part, one of converting raw optical variables to a second, more useful, form. i.e. vehicle state variables described in the inertial frame of reference. The psychophysical point of view described above assumes that no conversion is necessary, and that the human operates within an optically defined frame of reference. As a result, the control problem is one of selecting the most useful optical variables for specific control tasks and no frame-of-reference transformation is necessary. However, the describing function approach is more compatible with psychophysical investigations as it provides a useful tool for evaluating the optical variables that are correlated most highly with control behavior.

NASA RESEARCH PROGRAM

At NASA-Ames we have initiated a research program to understand and model how humans control flight through visual cues. One element has focused upon the value of formulating manual flight control as a problem in selecting and directly controlling optical variables. Toward this end, we have begun by examining flight control strategies in a minimally complex simulation of a visual hover task (see Figure 1.) This task (described more fully in two other reports [5], [6]) uses a simplified vehicle model with only three translational degrees of freedom: longitudinal (fore/aft), lateral (left/right), and vertical (up/down). No rotational motions are simulated. The human operator is given control over only vertical velocity, and told to maintain a constant altitude over the simulated ground plane.

The human operator's task is to use control stick motion to maintain a reference altitude over a grid plane in the presence of longitudinal, lateral, and vertical disturbances. Figure 2 shows the geometric pattern that the operator sees through the "windscreen" of the simulator. This represents what a pilot might see looking out of the window of an aircraft. It shows: (1) a set of ground "meridian" lines that are parallel to the forward gaze direction and fan out from the vanishing point on the horizon; and (2) a set of ground "latitude" lines that cross the field of view horizontally. No other information (i.e. flight instruments) is provided. This perspective view of the grid plane provides a host of potentially useful features or cues that relate in some analytical way to vehicle state variables [x (longitudinal), y (lateral), and h (vertical.)]

Three grid-plane patterns were studied: (1) a wire frame made of lines parallel to the forward gaze direction (meridian grid); (2) a wire frame made of lines orthogonal to the forward gaze direction (latitude grid); or (3) a wire frame made of both orthogonal and parallel grids (square grid). In addition a random terrain structure composed of irregular colored polygons was presented. This condition included all of the optical information available in the square grid, but in a stochastic fashion.

Performance was very good and nearly identical for trials with the square and latitude grids and with the terrain structure. Performance was poor with the meridian grid. For the square and latitude grids and the terrain structure, there was power in the stick output (stick motion) associated with the x disturbances as well as with the h disturbances; operators selected and regulated some optical variable(s) that produced stick inputs associated with changes in longitudinal craft position x in addition to craft altitude h. In control terminology, the stick motion showed the presence of an (undesirable) crossfeed from the craft's longitudinal motion, suggesting the choice of optical variable(s) that varied both with altitude and longitudinal motion.

An examination of the optical variables present in the three grid conditions revealed several cues which unambiguously relate to vehicle altitude alone (i.e. are invariant over changes in x and y). The operator could have used any of the following cues, which vary with altitude alone:

Cue (1)   The distance between any two points where the meridians intersect the bottom of the window (e.g. distance between A and B in Figure 2)

109

Cue (2) The number of image latitude lines on the window between any two window locations (e.g. three between M and N in Figure 2).

Cue (3) The number of image meridian intersections with the bottom of the window (e.g. the five intersections in Figure 2).

Since performance was generally better when grids with latitude lines were explicitly (i.e. the square and latitude grids) and implicitly present (the terrain structure), this might suggest that operators tried to focus on Cue 2; this is the only cue that depends solely upon latitude lines. However, the presence of the significant crossfeed of the longitudinal disturbance into control motion suggests that the operators must have used a mixed cue that reflected both vertical and longitudinal motion. One such cue is:

Cue (4) The visual optical depression angle of a ground latitude line below the horizon. (e.g.in Figure 2 this this is the visual angle, alpha, subtended by the distance, D, of the latitude line image below the horizon)

However, this observation is not a sufficient test of whether or not this cue was used for this hover task. One should be able to identify the specific reference depression angle that accounts for the observed time history of the stick motion and the corresponding performance data. Use of a given reference depression angle, alpha, implies that: (1) the describing functions relating altitude (h) and longitudinal position (x) to stick motion have the same shape, and (2) the ratio of low frequency h and x gains equal the tangent of alpha.

This technique was used to determine alpha and the corresponding describing functions. The stick response of this model closely follows the data. Operator control response and performance can also be described by using an optimal control formulation. Since this is a simple task, the internal model of the optimal control formulation would assume a representation which includes, at least, the two vehicle and two disturbance state variables associated with x and h. The presence of x crossfeed in the stick motion can only be accounted for by choosing a cost function that includes both x and h in addition to the control stick motion. However, it does not seem reasonable or parsimonious to assume that a person has an independent estimate of h but does not use it affect control.

CONCLUSION

Our initial results show that the use of control engineering modeling techniques, together with a psychophysical analysis of information in the perspective scene, holds promise for capturing the manual control strategies used during visual flight. It is important that we analyze behavior in this way before concluding that the description of visual flight control will be a simple modification of previous models. It is premature to conclude that, simply because humans can get around in a three-dimensional world in a very capable fashion, that they do this by extracting these dimensions and controlling their vehicles with respect to that three-dimensional frame of reference. For the purpose of control they may remain within the optical frame of reference.

REFERENCES

1. A.V. Phatak, M.S. Karmali and E.J. Hartzell, "Development of a pilot model for helicopter visual flight task segments," in AIAA Proceedings on Flight Testing to Identify Pilot workload and Pilot Dynamics, Edwards Air Force Base, CA. AFFTC-TR-82-5, Edited by Michael Frazier and Robert B. Crombie, 1982.

2. J.J. Gibson, The Senses Considered as Perceptual Systems, Houghton Mifflin Co., Boston, 1966.

3. D.T. McRuer and H.R. Jex, "A review of quasi-linear pilot models," IEEE Trans. on Human Factors and Electronics, HFE-8, 1967.

4. D.L. Kleinman, S. Baron, and W.H. Levison, "A control theoretic approach to manned-vehicle systems analysis," IEEE Trans. on Auto Control, AC-16, 1971.

5. W.W. Johnson and A.V. Phatak, "Optical variables and control strategy used in a visual hover task," in Proceedings of IEEE Conference on Systems, Man & Cybernetics, Cambridge, MA: November 1989.

6. W.W. Johnson, C.T. Bennett, K. O'Donnell and A.V. Phatak, "Optical variables useful in the active control of altitude," in Proceedings of the 23rd Annual Conference on Manual Control, Cambridge, MA: June, 1988.
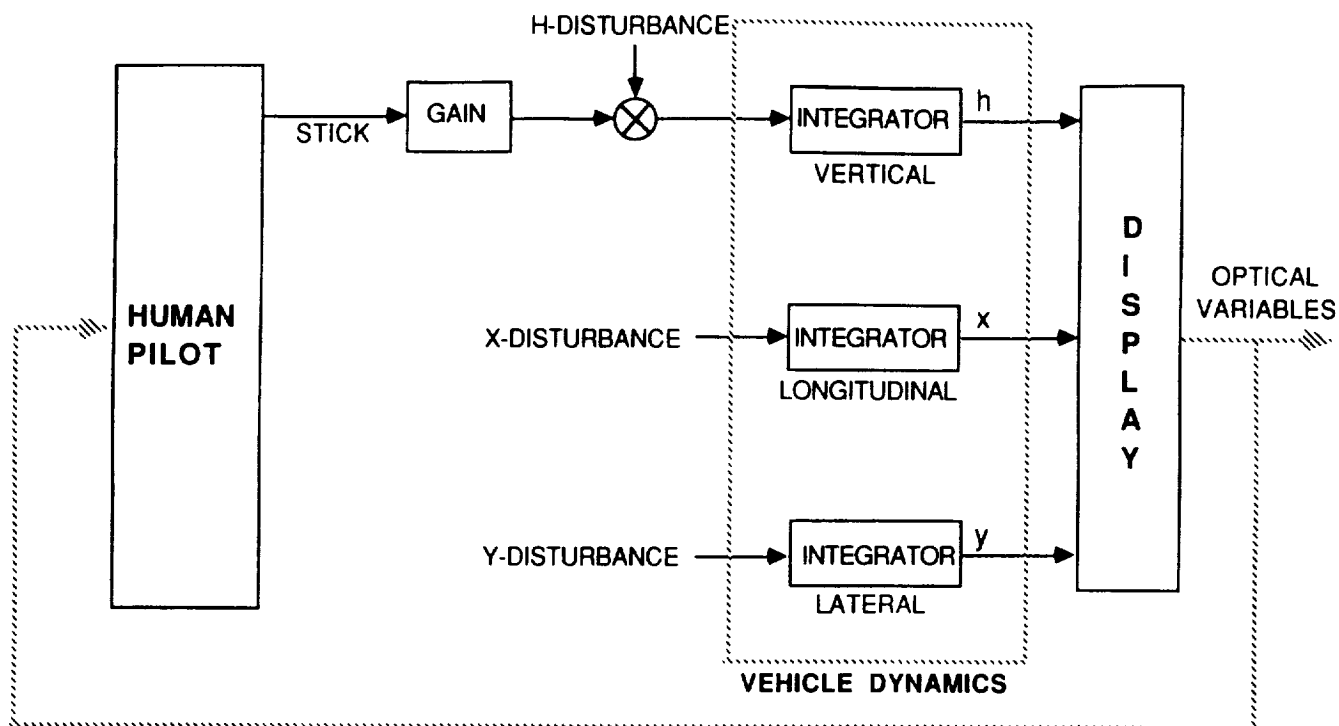
Figure 1. Functional block diagram of visual hover task.
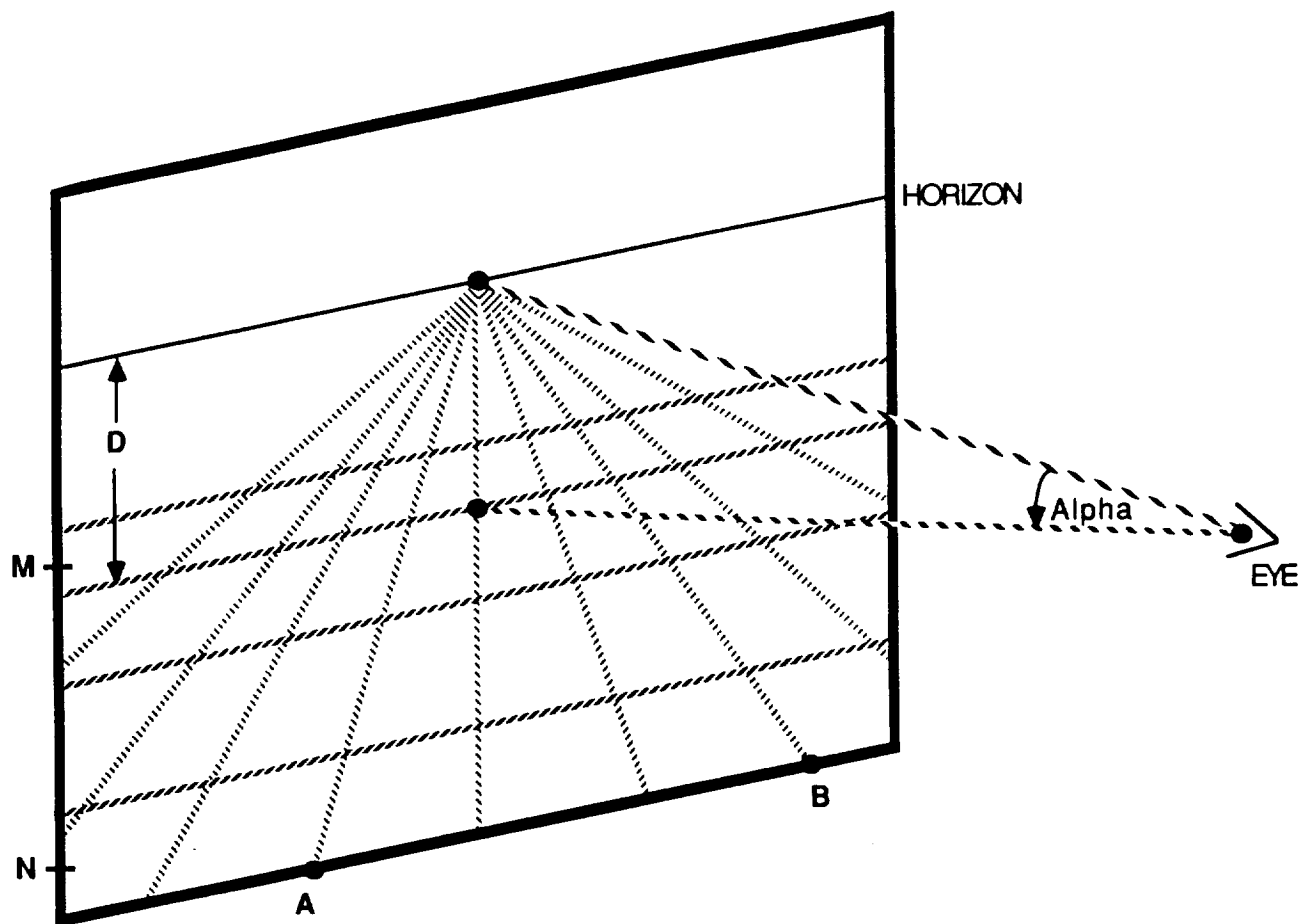
Figure 2. Out-of-the-window view from simulated vehicle cockpit.

SIMPLE CONTROL-THEORETIC MODELS OF HUMAN STEERING ACTIVITY IN VISUALLY GUIDED VEHICLE CONTROL

R. A. Hess
University of California
Davis, California

ABSTRACT

A simple control theoretic model of human steering or control activity in the lateral-directional control of vehicles such as automobiles and rotorcraft is discussed. The term "control theoretic" is used to emphasize the fact that the model is derived from a consideration of well-known control system design principles as opposed to psychological theories regarding egomotion, etc. The model is employed to emphasize the "closed-loop" nature of tasks involving the visually guided control of vehicles upon, or in close proximity to, the earth and to hypothesize how changes in vehicle dynamics can significantly alter the nature of the visual cues which a human might use in such tasks.

INTRODUCTION

The research to be briefly described stems from the author's participation in the Summer 1989 Workshop on the visually Guided Control of Movement, sponsored by the NASA Rotorcraft Human Factor Research Branch at NASA Ames Research Center. The approach to the Workshop theme discussed here is based almost entirely upon the human modeling paradigm which had its genesis in the work of feedback control engineers during, and immediately after, WWII [1]. The idea then, as now, was to compare the control behavior of the human to that of inanimate automatic feedback devices. The intervening 45 years has seen the discipline of manual control mature to the point that human performance, and to some extent, workload, can be predicted in certain well-defined control tasks with an accuracy sufficient for many problems of engineering design [2]. Based upon discussions at the Summer Workshop, the prevailing opinion among many psychologists is that the control theory paradigm has little more to tell us regarding human interaction with dynamic systems. This opinion may be premature.

A CONTROL THEORETIC MODEL FOR DRIVER STEERING BEHAVIOR

Automobile driving, or more appropriately, automobile steering, offers one of the simplest tasks involving human control of vehicle movement. The task is all the more attractive for discussion since it is one in which almost all humans above the age of sixteen participate daily. Figure 1 shows the steering task geometry involved in constant speed lane-keeping on a curving road. The variables

$y_V(t)$ and $y_R(t)$ represent vehicle and roadway lateral coordinates, respectively, and $\psi_V(t)$ and $\psi_R(t)$ represent vehicle and roadway heading, respectively.

A relatively simple control theoretic model for driver steering behavior can be offered as shown in Fig. 2 [3]. Space does not permit anything but a cursory description of this model. The interested reader is referred to [3]. Basically, the model is composed of high- and low-frequency compensation elements, defined by the transfer functions shown in Fig. 2. The high-frequency compensation is based upon a "structural model" of the human operator in which the compensation is achieved through proprioceptive, rather than visual, cues [4]. The low frequency compensation, denoted as G_C, is achieved through a simple visual guidance cue to be described shortly. It should be emphasized that, although nine parameters appear in the high-frequency compensation, all can be chosen based upon the vehicle transfer function $\dot{y}_V(s) / e_A(s)$ [3] and the dictates of the classical "crossover" model of manual control theory [5].

For the automobile steering task, feedback system design considerations dictate the form of $G_C(s)$ to be:

$$G_C(s) = u(s) / e_A(s) = K_Y(s) + (1 / T_3) \tag{1}$$

In the time domain, this transfer function translates into [6]:

$$\begin{aligned}
u(t) &= K_Y \left[\dot{e}_A(t) + (1 / T_3) e_A(t) \right] \\
&= K_Y \left[(\dot{y}_R(t) - \dot{y}_V(t)) + (1 / T_3)(y_R(t) - y_V(t)) \right] \\
&= K_Y \left[u_0(\psi_R(t) - \psi_V(t)) + (1 / T_3) y_E(t) \right] \\
&= K_Y u_0 \left[\psi_E(t) - \tan(\psi_1(t)) \right] \\
&\approx K_Y u_0 \left[\psi_E(t) + (\psi_1(t)) \right] \\
&\approx K_Y u_0 \psi_U(t)
\end{aligned} \tag{2}$$

where u_0 is the vehicle speed (assumed constant, here).

The last of Eqs. 2 is interpreted in Fig. 3. The variable, us, in the driver model of Fig. 2 is synonymous with the angle between the vehicle x-axis, x_B, and the line-of-sight to an "aim point" on the tangent to the roadway, a distance $u_0 T_3$ ahead of the vehicle. For most driving tasks, $t_3 \approx 3$ sec.

Using the driver model just described, very close agreement has been found between model responses and those obtained in driver simulation studies for a lane-keeping task on the curving roadway of Fig. 4 [3]. There is, of course, no psychological basis for the visual guidance cue just hypothesized. It may, in fact, not be a valid description for the actual visual field cue to cues used by the driver. However, the actual cues must, in a control theoretic sense, be equivalent to the cue just described.

116

A CONTROL THEORETIC MODEL FOR ROTORCRAFT NOE FLIGHT

Let us now consider that the vehicle shown in Fig. 1 is a rotorcraft in a nap of the Earth (NOE) mission in which the pilot is attempting to follow a groundtrack identical to the roadway of Fig. 4,m at the speed, u_o. As in the case of the automobile, we will consider only lateral-directional motion. Even so, the rotorcraft exhibits an additional degree of freedom, namely vehicle roll attitude ϕ. Now the same control theoretic model described in the preceding section for the high-frequency compensation can be applied to this problem, albeit with slightly different parameter values. Indeed, the same task variables and geometry as depicted in Fig. 1 are still valid. However, the fact that the vehicle dynamics have changed has a significant effect upon the form of $G_C(s)$ in the model. It can be shown that, in the case of the rotorcraft, the variable, u, is now given by:

$$u(t) = K_Y u_o \dot{\psi}_U(t) \tag{3}$$

thus, the time rate of change of the angle $\psi_U(t)$, or the angular velocity of the aim point line-of-sight is the visual guidance cue which can be hypothesized to be used by the pilot. Once again, there is no psychological basis for this cue, nonetheless, in a control theoretic sense, an equivalent cue or cues must be used by the pilot in this task.

CLOSURE

A simple control theoretic model of human steering behavior in a pair of vehicle control tasks with identical task descriptions has led to two different types of visual cues being hypothesized as central to successful task completion. The purposes of this admittedly rather crude study was to emphasize the fact that different vehicle dynamics can significantly alter the nature of the visual cues which a human might use in completing the task. This suggests that a study of the visually guided control of movement cannot neglect the fundamental feedback structure which permits such activity.

117

REFERENCES

[1]   Weiner, N., *Cybernetics*, Wiley, New York, 1948.

[2]   Hess, R. A., "Feedback Control Models" in *Handbook of Human Factors*, Ed: G. Salvendy, Wiley, New York, 1987, pp. 1212-1242.

[3]   Hess, R. A., and Modjtahedzadeh, "A Control Theoretic Model of Driver Steering Behavior," *IEEE Control Systems Magazine*, to appear.

[4]   Hess, R. A., "A Model-Based Theory for Analyzing Human Control Behavior," in *Advances in Man-Machine Systems Research*, Vol. 2, Ed: W. G. Rouse, JAI Press, London, 1985, pp. 129-175.

[5]   McRuer, D. T., and Krendel, A., "Mathematical Models of Human Pilot Behavior," AGARDograph No. 188, Jan. 1974.

[6]   Ogata, K., *Modern Control Engineering*, Prentice-Hall, Englewood Cliffs, NJ, 1970.
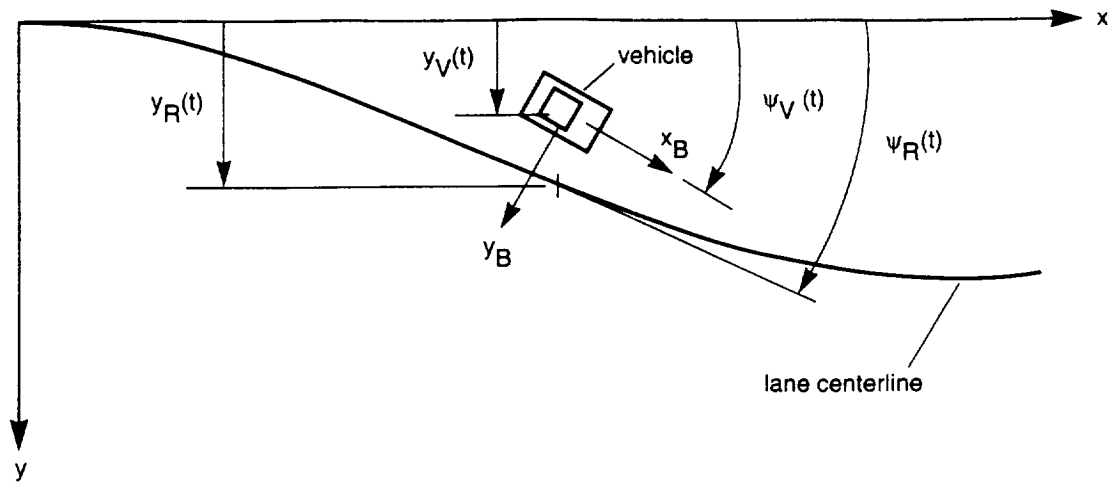
Figure 1. Steering task geometry.
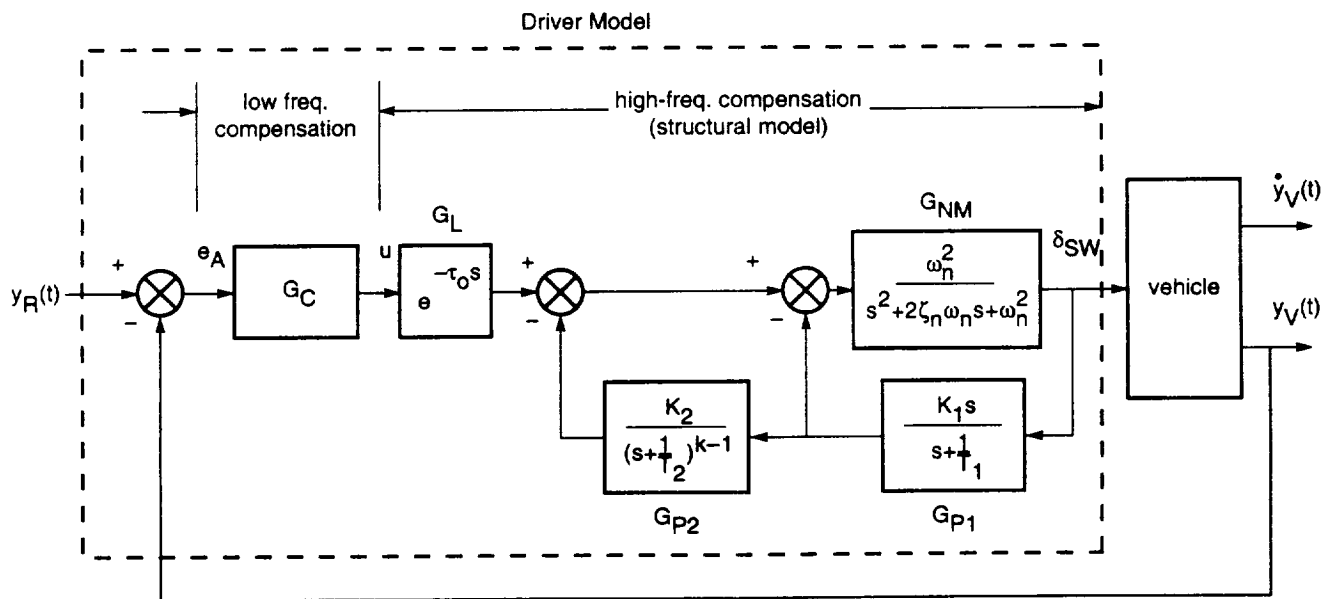
Figure 2. The driver/vehicle model.
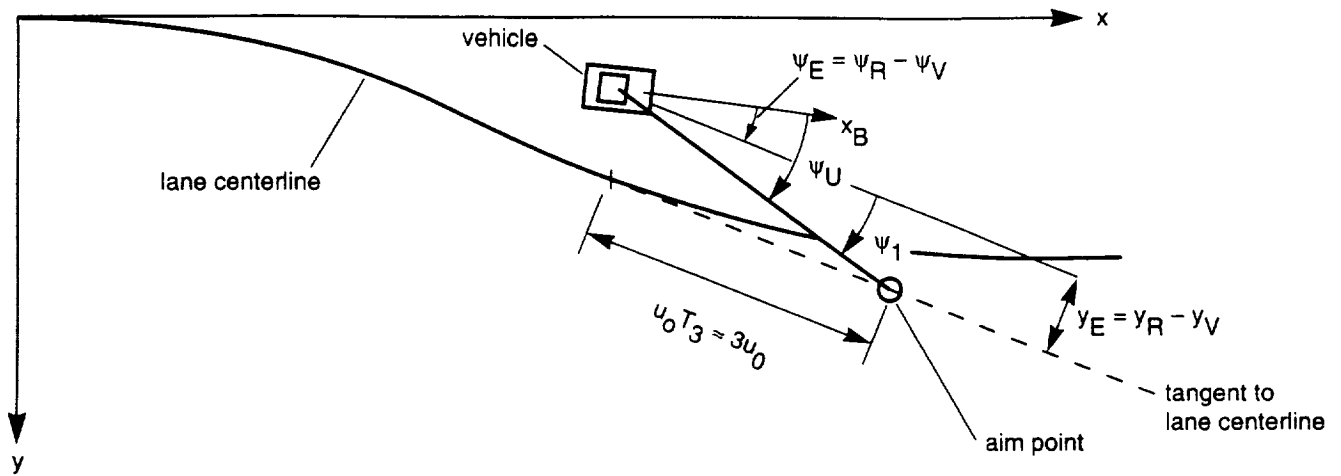
Figure 3. A visual guidance cue for the driving task.
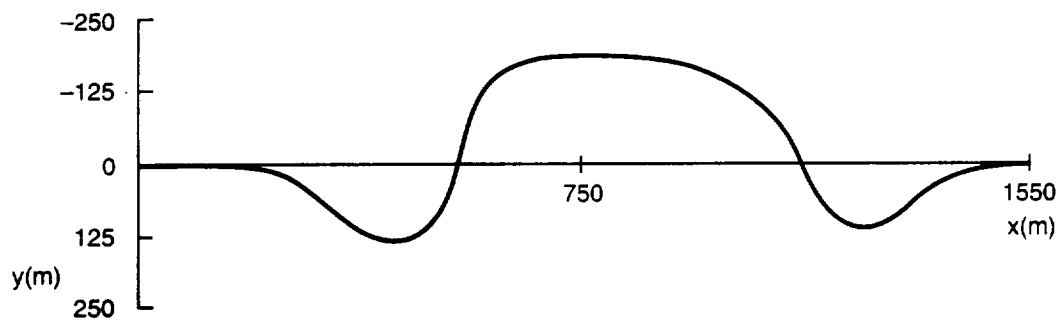
Figure 4. Curving roadway used in the driver simulation.

CONTROL WITH AN EYE FOR PERCEPTION: PRECURSORS TO AN ACTIVE PSYCHOPHYSICS

John M. Flach
Wright State University
Dayton, Ohio

ABSTRACT

The perception-action cycle is viewed within the context of research in manual control. A portrait of a perception-action system is derived from the primitives of control theory in order to evaluate the promise of this perspective for what Warren and McMillan (1984) have termed "Active Psychophysics." That is, a study of human performance that does justice to the intimate coupling between perception and action.

INTRODUCTION

Are there important differences between a human actively involved in accomplishing a goal directed activity and a human passively monitoring and making judgements about stimulation imposed from without? In the active mode the subject has control over stimulation. In the passive mode stimulation is controlled by an entity (generally the experimenter) other than the subject. These two modes may be different in terms of the control of attention; in terms of the kinds of information available; in terms of sensitivity to information; and are certainly different in terms of the kinds of activities required of the subject. Certainly Gibson's early studies with touch suggest that active and passive modes are fundamentally different in the kinds of information picked up by the actor/observer (Gibson, 1962). Stappers (1989) has recently shown that active control enhances visual form recognition. Also, research on the effects of automation on the performance of human-machine systems (out-of-the-loop syndrome) suggests that there are fundamental differences between systems where the human functions as a controller compared to systems where the human functions as a monitor (e.g. See Wickens, 1984, P.492). To the extent that the actor and the observer are different, care must be taken with how researchers generalize the results of experimentation. The domination of passive modes of interaction in psychological research (even in ecological research which is based on the concept of the perception-action cycle) may lead to inappropriate generalizations. For this reason a number of people (e.g. Warren & McMillan, 1984) have pointed out the need for research paradigms that permit subjects to actively control stimulation in pursuit of goals. In this paper, a tutorial review of control theory will be presented as one framework within which an "active psychophysics" might be pursued.

INPUT AND OUTPUT

Figure 1 shows a black box representation of a human-environment system. There are two qualitatively different sources of input into this black box and a single output. These inputs and outputs are not single dimensional entities but instead should be considered multidimensional vectors. The distinction between Intention and Disturbance, as qualitatively different inputs to the black box is critical for understanding the behavior of control systems. However, this distinction is often obscured in the literature on manual control. The term input is sometimes used to refer to intention and sometimes to disturbance (Powers, 1978). In general, a good controller will minimize the match between disturbance and output and will maximize the match between intention and output. In other words, a controller will behave so as to accomplish intentions (goals) and will do so in spite of any external disturbances that might perturb the system. The prototypical example is a thermostat. A temperature is input as an intention and this temperature is attained and maintained in spite of external inputs (disturbances) arising as a function of outside temperatures.

A second qualitative distinction is important in characterizing the input signals (both intentions and disturbances). Inputs can be discrete or continuous. An example of a discrete input used in the study of human performance is the Fitts' Law paradigm (see Jagacinski, In Press for review). The appearance of the target is an intentional input in which the goal of the operator is changed instantaneously from one position (the home position) to a second position (the target position). Step tracking is another example in which discrete signals (instantaneous changes of position) are used as inputs. When step tracking is performed in a pursuit mode, as illustrated in Figure 2, then the input is an intention. When step tracking is performed in a compensatory mode, then the input is a disturbance. In discrete control paradigms, dependent measures that are often used include:

Reaction Time – the time from the input signal onset to the onset of the response to that signal. This is illustrated in Figure 2.

Movement Time – the time from the initiation of a response to the input signal to the completion of the response (e.g., target capture).

Accuracy – the match between intention and action (output) at the end of a response sequence.

Submovements – often the output resulting from a discrete input can be parsed into segments (e.g., submovements). Important measures include the number of submovements; the duration of individual submovements; the accuracy of individual submovements; the peak velocities; and the peak accelerations.

Continuous signals can also be used as input to the black box. Typically, the continuous signals used in manual control experiments are constructed as a sum of sine waves. There are two reasons for this choice. First, Fourier's Theorem shows that any periodic signal can be approximated as a sum of sine waves. Thus, sine waves are fundamental building blocks for constructing a wide range of signals. A second reason for using sine waves to construct signals is that for a linear servomechanism a sine wave input will result in a sine wave output at the same frequency, but changed in amplitude and phase. The pattern of amplitude and phase changes can be extremely useful for drawing

inferences about the nature of the black box (e.g., the transfer functions). Also, frequency can be used as a signature to differentiate the sensitivity of the black box to various kinds of inputs. The use of frequency signatures to differentiate sensitivity will be discussed further in a later section of the paper. When continuous signals are input as intentions, then the subject's task is called a pursuit tracking task. In this task the subject sees both a continuously changing target (e.g., a roadway) and a cursor representing her position with respect to the roadway. A good controller would be one that minimized deviations between her position and target position. When continuous signals are input as disturbances, then the subject's task is called a compensatory tracking task. Here the subject's goal is a fixed position (e.g., center of screen or constant altitude) and a disturbance (e.g., windgust) is input that drives the subjects away from their fixed goal. In pursuit tracking, subjects can see movements of the goal and movements of themselves with respect to that goal. In compensatory tracking, subjects see only their own movement with regard to the fixed goal. For research using continuous inputs the dependent variables typically used include:

RMS Error – this is the square root of the sum of squared deviations between cursor (ego or vehicle) position and the goal position (summed over samples) divided by the number of samples. This method of scoring results in a differential weighting of small and large errors.

Small errors contribute proportionally less to RMS error than do large deviations.

RMS Control and RMS Control Velocity – these measures are similar to RMS error. They are indexes of the amount of control activity.

Time-on-Target (TOT) – this is a measure of the proportion of time during a tracking trial that the subject is within the boundaries of the target.

Amplitude and Phase – the amplitude and phase are measured at each frequency of input. The ratio of amplitude in the output to amplitude in the input signal is termed gain. These measurements are important for characterizing the transfer function of the black box.

Remnant – the remnant is the output power at noninput frequencies. This is an index of the control variance that is not correlated with input signals.

NEGATIVE FEEDBACK CONTROL

A simple system that acts to attain and maintain an intention in spite of disturbances is a negative feedback system. Figure 3 shows a simple negative feedback device. The new ingredient that the negative feedback system introduces is error. This is the difference between the intention or goal and the current state of the system. A negative feedback system is driven by error. that is, when error is zero there is no action in this system. When error is non-zero this system will attempt to reduce the error. Whether or not the system is successful in reducing error will depend on the characteristics of G. Figure 3 shows a derivation of the relation between Intention, Disturbance, and Output as mediated through G. The equation relating these elements is:

$$[G/(1+G)]*\text{Intention} + [1/(1+G)]*\text{Disturbance} = \text{Output} \qquad (1)$$

Note from Equation 1 that if $\underline{G}$ is a simple multiplier then the greater the value of $\underline{G}$ (i.e., the higher the open loop gain) the closer will be the match between Output and Intention. The term that operates on Intention will go to 1 as $\underline{G}$ becomes large. The term that operates on Disturbances will go to 0 as $\underline{G}$ becomes large. Thus, as $\underline{G}$ becomes large Equation 1 will reduce to:

$$\text{Intention} = \text{Output} \qquad (2)$$

In nature $\underline{G}$ is never a simple multiplier. For all physical systems there will be a delay associated with $\underline{G}$. For control purposes it is not the absolute time associated with this delay but the time relative to the frequency of the signal. That is, the key dimension will be the proportion of a cycle that a signal is delayed. This is termed phase lag. If a signal is delayed by 180 degrees then the negative feedback system will result in a diverging error. Such a system is said to be unstable. For good control $\underline{G}$ should have high gain when the phase lag is less than 180 degrees. The higher the gain, the faster error will be reduced. $\underline{G}$ should have low gain, less than 1, as the phase lag approaches and exceeds 180 degrees. This relation between gain and time delay is illustrated in Figure 4, which is adapted from Jagacinski (1977). The graph shows three regions sluggish control, good control, and unstable control. If the time delay is small (small phase lag) and the gain is low then error will be reduced very slowly. An example of a sluggish response to a step input is shown in Figure 4. If the time delay is large and gain is high the error will not be reduced and in fact will become greater. This is the region of unstable control. Pilot induced oscillations in flight result from a pilot responding with two high a gain given the time delays associated with the system. An example of an unstable response to a step input is also shown in Figure 4. If gain is high and time delay is small or if gain is low when time delay is large then good tracking will result. Two examples of the response of a good tracker to a step input are illustrated in Figure 4. Note that as the time delay becomes greater the range of gains that will result in good tracking diminishes.

The relationship between gain and phase lag can also be illustrated using a Bode plot. The Bode plot shows open loop gain (in decibels) and phase lag (in degrees) plotted as a function of the log of frequency (in radians/sec). Figure 5 shows the pattern of gain and phase lag that would be obtained for a good controller. This pattern represents good control in that for those frequencies with phase lag less than 180 degrees gain is high. Thus, intentional signals at those frequencies will be followed closely in the output and disturbances at those frequencies will be filtered out (will not show up as output). In other words, errors will be eliminated quickly. For those frequencies with phase lags greater than 180 degrees gain is less than 1. Thus, the system will be stable. Intentional signals at those frequencies will not be followed in the output and disturbances at those frequencies will not be filtered out (they will be part of the output).

A key landmark in the Bode plot is the "crossover point," the point at which gain is equal to 1 (0 db). For the system to be stable the phase lag must be less than 180 degrees at that point. the distance of the phase lag from 180 degrees is called the phase margin of the system. A positive phase margin is required for stable control. The frequency of the crossover point indicates the bandwidth of the controller. Intentional signals at frequencies below the crossover point will be represented in the output. Intentional signals at frequencies above the crossover point will be filtered out (will be attenuated in the output).

A final point to be noted about negative feedback, closed-loop systems concerns the concept of time. The common sense notions of before and after do not apply. Errors do not precede actions which in turn precede feedback. Errors, action, and feedback are continuously available. In place of the common sense notion of time is the concept of phase. Action can be in-phase with feedback (perception) or out-of-phase. When in-phase the system will be stable. When sufficiently out-of-phase the system will be unstable.

MANUAL CONTROL

Manual control is the study of negative feedback control systems in which the loop is closed through a human operator. That is, the human operator is given a task or goal to accomplish this goal is accomplished by observing displays and manipulating controls. This situation is illustrated in Figure 6. Note that the G in the forward loop of Figure 4 has been replaced by two boxes in the forward loop of Figure 6. One box, labelled Controller, represents the transfer function for the human operator. The second box, labelled Plant, represents the transfer function for the physical system that the human is interacting with (e.g., dynamics of the helicopter). The central problem for a theory of manual control has been to build a model or theory of the human operator. Two approaches to modeling the human will be distinguished. One approach assumes that the human operator responds continuously to error. The second approach assumes that the human responds in a discrete fashion.

Continuous Control

Early researchers began with the assumption that the transfer function of the human operator would be invariant, independent of the plant dynamics. It was assumed, that once this transfer function was discovered it could be used to predict performance across a wide range of plant dynamics. McRuer and his colleagues (e.g., McRuer & Jex, 1967; McRuer & Krendel, 1974; McRuer & Weir, 1969) soon discovered that this definitely was not the case. As the dynamics of the plant changed, so to, did the describing function for the human operator. The invariant, as McRuer et al. discovered was not at the level of the human but was at the level of the total forward loop (human + plant) describing function. This invariant at the level of the human/plant combination was the basis for the classic "crossover" model. The key insight behind the crossover model is illustrated in Figure 7. The first column in Figure 1 shows Bode diagrams and transfer functions [using Laplace notation] for three simple plant dynamics. The second column in Figure 7 shows describing functions obtained for humans controlling each of the three dynamic plants. The final column shows the describing function for the human/plant combination. Note that the patterns in Column 3 are invariant and that they have the same form as the "good" controller illustrated in Figure 5. What was surprising to earlier researchers should be obvious in retrospect. The constraints on good stable performance operate at the level of the total forward loop (human + plant). To do the task the human must operate within those constraints and therefore must adapt to the plant dynamics in a way that is consistent with those constraints. Thus, the "crossover" model predicts that in the region of crossover the human plus the plant will approximate the transfer function shown in Column 3 of Figure 7.

In adjusting to the plant dynamics to both satisfy the demands to minimize RMS error and to satisfy the constraints for stability the human behaves like an optimal controller. This observation

125

was the basis for the "optimal control" model of the human operator (e.g. Baron & Kleinman, 1969; Kleinman, Baron, & Levison, 1970; Kleinman, Baron, & Levison, 1971). The optimal control model assumes that the human operator uses an internal (mental) model of the plant dynamics to estimate the current states of the system from delayed, noisy observations of display position and velocity. The human responses to these states are based on an optimal control law which chooses response gains that minimizes a linear combination of squared tracking error and squared control velocity. Thus, in a sense, the model assumes that the operator attempts to achieve minimum error with minimum effort. These responses are filtered through the limb dynamics and are contaminated by motor noise.

The optimal control model has been popular because there is a natural mapping from the elements of the model to the stages (encoding, estimation, decision, response) of the standard information processing model that has dominated modern psychology (See Pew & Baron, 1978). The optimal control model also provides a better fit over a wider range of frequencies to human performance data than does the crossover model. However, to do so it requires a greater number of parameters.

The crossover model and the optimal control model both assume that the human responds in a continuous, proportional (linear) fashion to error and error velocity. However, there is much evidence that the human is not linear (e.g. see Knoop, 1978). For example, there is the presence of remnant in the human control response. Remnant is power at output frequencies not present in the input. As noted in an earlier section, a linear system would only have output at the input frequencies. The optimal control model accounts for the remnant by assuming the presence of broad band white noise injected by human perceptual and motor processes. The non-white shape of the measured remnant is thought to reflect the dynamics of the humans' perceptual and motor processes. Others have argued that the remnant arises, at least in part, due to the discrete, nonlinear nature of the human transfer function.

Discrete Control

In discussing discrete control models of the human operator three classes of models will be presented—synchronous discrete controllers, asynchronous discrete controllers, and hierarchical controllers.

Bekey (1962) lists a number of studies that have found evidence of a "psychological refractory" period when a human is required to respond to discrete stimuli spaced by less than about 0.5 seconds (Hick, 1948; Welford, 1952; Davies, 1957). One inference that might be drawn from this finding is that the human "acts on discrete samples of information from the external world." Figure 8, adapted from Bekey (1962) gives examples of two synchronous discrete controllers. These controllers act on discrete observations taken at a fixed frequency. A synchronous sampler with a 0-order hold responds as a function of the position observed at each sample. The synchronous sampler with a 1st-order hold responds as a function of the position and velocity observed at each sample. Three important attributes of synchronous discrete controllers noted by Bekey (1962) are:

(1)   Changes in the input cannot have any effect until the next sampling instant occurs.

(2)   The presence of the sampler limits the frequencies which can be reconstructed at its output to those not exceeding one-half the sampling frequency.

(3)   The action of the sampler generates harmonics in the output which extend over the entire frequency spectrum, even when the input is band limited. (Bekey, p. 45-46.)

The last attribute provides an alternative explanation for the remnant power routinely observed in human tracking data.

A synchronous discrete controller responds at a fixed frequency. An asynchronous controller responds at irregular intervals. Angel & Bekey (1968) have proposed a finite-state model for manual control that behaves asynchronously. The logic of the finite-state controller is illustrated in Figure 9. Inputs to this controller are coarsely quantized with regard to threshold boundaries on position and velocity. These boundaries are the dashed lines in Figure 9a. These quantized inputs are responded to with simple force time programs which are shown in each region of state space. For example a large position error with low velocity evokes a large amplitude bang-bang response. This type of model has great intuitive appeal for modeling human control of second-order control systems, where there is evidence that humans exhibit bang-bang control (Young & Meiry, 1965). This nonlinear style of control provides still another alternative explanation for remnant.

Costello (1968) proposed a model of the human tracker using a hierarchical control model. Costello's model is illustrated in Figure 9b. Costello proposed two modes of control. He proposed that the human controller responded to small errors and error velocities in a manner consistent with the crossover model. This is the central region of the state space identified with the constant coefficient mode. To large errors, the model predicts that the human will respond in a time optimal bang-bang fashion. Costello called this the surge mode. Jagacinski, Plamondon, and Miller (1987) have also employed a multi-level style of modeling in which a number of low level motion generators (Herding mode, predictive mode, close following mode, fast acquisition mode) are combined with finite state logic to model human performance in capturing evasive, moving targets.

SUMMARY

The continuous control models have dominated much of the work on manual control. These models have been useful tools for evaluating human control systems and for making predictions about stability of these systems. They have particularly been widely used for studying vehicular control. However, it is clear that some of the assumptions made by these models must be questioned. One must wonder whether the practical utility and success of these models has retarded scientific progress in understanding human control.

There is one intervening variable that should be considered when choosing between the linear, proportional control models (i.e., crossover, optimal control, synchronous controller) and the nonlinear, discrete control models (i.e. asynchronous or hierarchical controllers). That is the time lag of the physical system being controlled. The linear, proportional control models work well for systems that have small time lags (e.g., high performance aircraft). However, these types of models are totally

inadequate for systems with long time lags such as thermodynamic systems (see Crossman & Cooke, 1974). For slow responding systems it is clear that humans respond in a discrete, nonproportional fashion.

This has been a very brief and selective review of some of the models that have been proposed for the human controller. For the most part, the research that has inspired these models has employed simple laboratory tracking tasks using moving cursors on CRT displays. In this kind of task the error signal is clearly defined and thus the perceptual problems have not generated very interesting problems. It remains for an ecological psychology to study control behavior with less well defined error displays (e.g., optical flow fields). This review is presented here because as the perceptual problems are addressed, our ability to draw correct inferences about perception will depend on our use of informed assumptions about action.

Closing the Loop Through the Optic Array

"...instead of searching for mechanisms in the environment that turn organisms into trivial machines, we have to find the mechanism within the organisms that enable them to turn their environment into a trivial machine." (von Foerster, 1984, p. 171)

The laboratory tracking task, in one sense, is a task that turns humans into a trivial machine (e.g. a simple gain, integrator, or differentiators). The error signal and the goal of the operator are "trivial" relative to the signals by which humans control their own locomotion. The problem in more natural environments is not simply to generate the appropriate control law, but to extract from the "booming, buzzing confusion" the information that specify the goals and the error with respect to those goals. Gavan Lintern (personal communication) has observed that, when learning to fly, controlling the airplane (getting it where you wanted it) was not the problem. The problem was knowing where you wanted to be. That is knowing what the correct glideslope looked like. A critical aspect of the organism turning its environment into a trivial machine may be an ability to pick-up information about regularities in the environment. Thus, it is the tuning to invariants in perceptual arrays that allows the "booming, buzzing confusion" to be managed. How information (i.e. invariants, constraints, or structure) in the optic array supports action has been a central question for ecological psychology ever since Gibson, Olum, and Rosenblatt's (1955) classic analysis of parallax and perspective during aircraft landings. However, in asking questions about the pick-up of information from optic arrays there is little evidence of a commitment to "active vision." Many of the studies of information pick-up have employed passive psychophysical methodologies (e.g., Warren, 1976; Owen, Warren, Jensen, Mangold, and Hettinger, 1981; Cutting, 1986; Anderson and Braunstein, 1985; Warren, Morris and Kalish, 1988; Larish and Flach, in press). Not only have our experiments employed passive tasks, but Stappers, Smets, and Overbeeke (1989) have argued that our conceptualizations of the flow field and of the information within it have been founded on the image of a passively translating, disembodied eye. They argue that many of the classic ambiguities disappear when one considers the information in optic flow fields generated by bouncing eyes locomoting over a surface of support. Stappers, et al. note that formal accounts of optic flow (e.g. Longuet-Higgins and Prazdny, 1980; Koenderink, 1986) "neglect the fact that the optic flow is largely brought about by the actions of the observer, and for just this reason it can be relative to the observer's effectivities: the observer's actions scale the information he samples."

128

The Performatory Loop

Figure 10 illustrates an initial framework for asking questions about the perception-action cycle where the loop is closed through an optic array. In this framework, the human observer is given an implicit (e.g. maintain stable posture) or explicit (e.g. maintain a constant altitude) goal. Control activity is then measured as a function of manipulations of the optic array (e.g. front vs. side view, lamellar vs. radial flow, parallel vs. perpendicular texture). A number of studies have begun to appear that have been framed in this manner. Stoffregen (1985) and Andersen and Dyer (1989) have used postural regulation as a control problem within which to study optic flow. Owen and Warren (1987) report research that examined control responses to discrete changes in acceleration and to ramp changes in altitude in order to identify the optical information that specifies egospeed and altitude. Warren (1988) review a series of studies that have examined altitude control with a continuous, sum-of-sines disturbance. Within this framework, Warren has varied the nature of the optic array (e.g. presence of perspective roadway) and the nature of the task (e.g. altitude maintenance, or fly as low as possible) in order to isolate the functional optical information for altitude. Johnson, Bennett, O'Donnell, and Phatak (1988) have also used an altitude regulation task to examine the utility of alternative structures in the optic array.

The Johnson et al. paper is particularly useful for illustrating the promise of control theoretic methodologies for supporting inferences about perception and action. In order to highlight the logic of the control methodologies the details of the experiment will be greatly simplified. Johnson et al. were interested in comparing the relative efficacy of two sources of optical information about altitude—splay angle and optical density. To address this question displays were chosen which isolate the two sources of information. These are shown in Figure 11a. Texture parallel to the direction of travel contains splay. Texture perpendicular to the direction of travel contains optical density but no spay. Square texture combines both splay and optical density. Crossed with the type of display were three types of disturbance. A horizontal disturbance (altitude) affected both parallel (splay) and perpendicular (optical density) texture. A fore-aft disturbance (headwind) affected only perpendicular texture. Finally, a lateral (side-to-side) disturbance affected only parallel texture. The three disturbances were constructed from sine waves so that the bandwidths of the disturbances were similar, but so that the frequencies were specific to a disturbance (no shared harmonics). This is illustrated in Figure 11b. Frequency can now be used as a signature to identify the control activity specific to optical features. Johnson et al. found better control of altitude with perpendicular texture. They also found that there was more altitude control resulting from the fore-aft disturbance (seen only in perpendicular texture), than from the lateral disturbance. This provides strong evidence that for the hover task studied, perpendicular texture provided a powerful source of information, guiding altitude control behavior whether it was specific to altitude or not.

Exploratory Behavior

The framework in Figure 10 represents an advance over passive psychophysics. However, experiments designed within that framework, still constrain the human to behave as a rather simple machine (servomechanism). In the framework of Figure 10 behavior arises only as a function of error with respect to performatory goals. However, humans act, not only to accomplish performatory goals, but also, humans act to pick-up information. Humans actively explore the environment. This exploratory mode of behavior is intimately coupled with performatory modes of behavior.

129

Information picked-up through exploratory activity will often support performatory activity. Also, performatory activity will itself result in the pick-up of information. An important challenge for an active psychophysics will be the study of the coupling of performance with exploration. Experimental paradigms must include tasks that allow or even encourage exploration. Active psychophysics must explore measurement and analysis techniques for parsing exploratory and performatory activities; or must discover meaningful higher-order parameters for gauging the interaction of exploratory and performatory modes.

One basis for parsing exploratory and performatory activities might be the distinction between correlated and uncorrelated power resulting from frequency analyses of control behaviors. For systems with small time constants and for well trained operators it might be expected that performatory activities will be closely linked to the "driving function" (i.e., the changing goal or the disturbance that perturbs the system from a fixed goal). Thus, performatory activity will be task driven. Exploratory activity, however, originates with the operator. This will likely be uncorrelated with the driving function and therefore, will appear as remnant. As we have seen earlier in this paper exploratory activity will probably not be the only source of remnant. Other sources that have been considered include perceptual/motor noise, discrete response strategies, nonlinearities, and uncorrelated optical activity. Remnant appears to be rich in information about the human operator. In fact, it could be argued that most of the psychology resides in the remnant. Whereas the correlated power carries little information about the operator, informing us, rather about the task.

Higher order parameters for gauging the interaction of performatory and exploratory modes might be stability and bandwidth. As operators discover more effective ways to pick-up information, this should be reflected in either larger stability margins or in greater bandwidths.

Questions about remnant may be the only avenue for addressing the performatory/exploratory distinction within the experimental framework shown in Figure 10. In this framework there is only a single response channel for both exploratory and performatory activities. Frequency analysis is a useful tool for partitioning different signals within a single channel. It may be easier to study performatory/exploratory interactions if our experimental framework is expanded to permit a second channel of activity. A natural choice for this second channel of activity would be eye movements as shown in Figure 12.

While it is not impossible to imagine situations where eye movements can have a performatory function (e.g., social interactions), in many natural task situations eye movements are purely exploratory. That is, they have no direct effect on error with respect to performatory goals. The indirect effects, however, may be great in terms of the information pick-up that the eye movements mediate. For this reason, the study of eye movements must be a critical element within an active psychophysics.

When the possibility of eye movements is introduced an important theoretical question must be addressed. This involves the question of whether information is specific to an ambient optic array or to the retinal array. For example, the focus of expansion (Gibson, 1947; 1950; 1958; see also Warren, Morris, and Kalish, 1988) is an invariant that specifies the direction of locomotion which has been defined relative to the ambient optic array. That is, the focus of expansion is a pattern within optic flow that arises as a consequence of a moving observation point. This pattern is a

consequence of ecological optics—the properties of light. It is independent of the nature of a sensory mechanism (e.g. simple vs multifaceted lens) and is independent of the viewport (i.e., where the organism is looking). On the other hand, Cutting's (1986) differential motion parallax has been proposed as an alternative invariant specifying direction of locomotion that has been defined with respect to the retinal array. That is, the invariant relations of differential motion parallax are specific to a viewpoint. They depend on a particular point of fixation.

I assert that both the ambient optic array and the retinal array descriptions have an important place in an active psychophysics. The world (including the observer) structures the ambient array. The structure in the ambient array is information about the world and the observer. This structure is present at a station point and in the relations between station points. Pick-up of information requires first a transducer sensitive to the energy that carries the structure. Second, pick-up depends upon activity (sampling). What information is picked up depends on the activity of the observer? A stationary observer can pick up only the information at a single station point. This is an extremely impoverished view. A moving observer has access to information from multiple station points and has access to the information in the relations across station points. Note that no information about environmental layout is created by movement. The information exists whether the observer moves or not. Movement simply makes the information available. Also note, that a particular movement only provides access to the information at the station points sampled and in the relations across those station points. Some ways of acting will reveal different information than others. Therefore, some ways of acting will be more effective for certain tasks, because the information made available will be more appropriate.

An important challenge for an active psychophysics will be to provide a framework for evaluating the effectiveness of sampling behavior. The challenge is not to provide an absolute metric for effectiveness, because effectiveness can only be measured relative to a task, but to provide a collection of methodologies for asking questions and drawing inferences with regard to sampling behavior. Thus, it is meaningful and important to ask the following question: For a given pattern of sampling behavior what information is in principle available to the actor/observer? This is where the retinal array becomes important. The retinal array is one kind of record of the information made available by a particular pattern of sampling behavior.

Mathematical descriptions of the retinal array can be very useful for generating hypotheses about what subset of the information from the ambient array is preserved over a particular set of samples from that array. However, it is important to note that there is an asymmetry in the logic of mathematical descriptions of both the ambient field and the retinal field. If an invariant mathematical relationship can be demonstrated between structure in the ambient array (or structure on the retina) and properties of the world (including observer) then this is proof that information is present. However, failure to discover a mathematical relationship does not prove that there is no structure. In this sense, no particular form of mathematical representation is privileged.

An active psychophysics must appreciate the importance of mathematical analyses of the ambient array and of the retinal array. However, it should never be constrained by these analyses. These mathematical analyses will help us to discover what are interesting questions to ask. However, the answers can only come from observations of behavior.

For an active psychophysics to be complete, observations must be made in which the actor has unrestricted and independent control over performatory and exploratory modes of behavior. In all of the studies cited above that examined control through optic arrays, performatory and exploratory behavior were constrained so that the actor could only look where he was going. However, in most natural environments no such restriction is present. When given independent control of exploratory behavior, where do people look? Are some patterns of looking more or less effective than others? Do different pattens of looking result in qualitatively different styles of control? These are the kinds of questions that motivated Gibson's (1962) observations on active touch (see also Stappers, 1989). These kinds of questions must be central to an active psychophysics.

Adaptation and Learning

Adaptation and learning are obvious and important side effects of the interaction between performatory and exploratory modes of behavior. Exploratory activity results in the discovery of information. The more information available to the actor the greater will be the number of control strategies that are available. A wider range of control strategies will open the possibility for both greater precision of control and greater stability. Figure 13 shows the addition of "adaptive logic" to our growing diagram of a perception/action cycle. Behind this small box hides enough mysteries to support many careers in Psychology.

The signals entering the adaptive logic box are of the same general nature as the signals throughout the network. These signals are patterns of energy in space-time and these signals are operated on by the boxes in the diagram. However, the signals leaving the adaptive logic box are different. They represent operators that operate on the other boxes. For example, output from the adaptive logic box may result in a change of the transfer function between observation of error and action. This results in an interesting circularity or coupling. The patterns of energy in space-time (connections between boxes) are both operators and operands. So to, the embodied constraints represented as boxes are themselves operated on by the very signals upon which they operate. This kind of coupling between system and signal is also seen in neural nets and connectionist machines that tune to invariants in stimulation (see McClelland and Rumelhart (1986) for review).

Control theoretic technologies may not be the most useful tools for organizing our thinking with regard to this coupling of system and signal. Field descriptions such as those described by Kugler and Turvey (1987) may be more useful. However, as we explore new modes of description we should proceed armed with the intuitions of those who have gone before. McRuer, Allen, Weir, and Klein (1977) have proposed the Successive Organization of Perception (SOP) model as a tool for understanding how the control logic might change with learning. This model, shown in Figure 14, includes three modes of tracking. The compensatory mode has been discussed throughout this paper. In this mode the human acts like a servomechanism responding to error between intention and output. The compensatory mode would dominate for a naive operator. As the operator becomes experienced he begins to learn the dynamics of the plant being control. Thus, he can anticipate the response of the plant. This allows him to respond directly to intentions rather than to error. To the extent that his anticipations are incorrect the residual error will be reduced as a result of the inner compensatory loop. If the environment that specifies the intention behaves in a consistent way (e.g. a track composed of a single sine wave), then the observer may tune to these consistencies. In other words, the operator may learn the "rule" or "pattern" that governs the input. This will allow a response to the

higher order pattern and free the operator from the requirement to continuously monitor input or error. This mode has been called precognitive. For example, an operator tracking a cursor driven by a single sine wave, may synchronize his response to the periodic pattern. Thus, the operator could close his eyes and still maintain close tracking (at least for short periods). While one mode or the other may dominate, depending on the state of the operator (e.g. experience level) or the state of the task (e.g. regularity), all modes are expected to operate in concert complementing each other.

Important empirical work has also been done on adaptation in the context of manual control (e.g. Young, 1969; Wicken, 1984). This empirical work should be instructive to those pursuing an active psychophysics. The following challenge from Young (1969) signifies the need for an active psychophysics to organize our thinking with regard to adaptive control.

"...what is being offered to solve the manual control problems of tomorrow? What will be the "critical task" facing the astronaut entering the atmospheres of a strange planet, the captain of an SST, the pilot of a commercial airliner making an approach in zero-zero visibility, the VTOL pilot guiding his unstable vehicle to a downtown landing field, the submarine commander, or the engineer on a high speed transportation system? Will they be involved in compensatory tracking? Obviously not. They will be on board for the versatility, adaptability, and reliability they add to an automatic system. They will be expected to observe the environment and use "programmed adaptive control" to change plans. They will monitor instruments and repair malfunctioning components. They will control in parallel with the automatic system and take over in the event of failure. What is the extent of the theory for predicting man-machine behavior in these simulations? It is almost nil." (Young, 1969, p. 329)

CONCLUSIONS

"The world is as many ways as it can be truly described, seen, pictured, etc. and there is no such thing as the way the world is." Nelson Goodman (1968)

Figure 13 represents one way to picture a perception/action cycle. It is not the way to picture perception/action cycles. The representation is not a roadmap for the future. In fact, it could be argued that if the representation in Figure 13 is taken too literally, then it will severely constrain our thinking and will be an obstacle to future progress. If the representation in Figure 13 is useful it is as a map to the past. That is, as a link to the study of manual control. The research on manual control has much to offer to anyone interested in the coupling of perception and action. As a new active psychophysics is molded, its shape should not be constrained by the cybernetic hypotheses that guided much of the work in manual control. However, our vision of the future of active psychophysics will be much clearer if we stand on the shoulders of those who have gone before. The methodologies of manual control offer an important alternative to the passive methodologies that dominate current psychophysics. If these methodologies are applied with caution and restraint, the future of an active psychophysics will hold great promise. Alternatively, the challenges posed by an ecological approach to perception and action promise to rejuvenate an area of research that is being lulled to sleep reliving past successes.

ACKNOWLEDGMENT

REFERENCES

Anderson, G. J. and Braunstein, M. L. (1985). Induced self-motion in central vision. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 122-132.

Andersen, G. J. and Dyer, B. P. (1989). Spatial orientation from optic flow in the central visual field. *Perception and Psychophysics*, 45, 453-458.

Angel, E. S. and Bekey, G. A. (1968). Adaptive finite-state models of manual control systems. *IEEE Transactions on Man-Machine Systems*, March, 15-20.

Baron, S. and Kleinman, D. L. (1969). The human as an optimal controller and information processor. *IEEE Transactions on Man-Machine Systems*, MMS-10, 1, 9-17.

Bekey, G. A. (1962). The human operator as a sampled-data system. *IRE Transactions on Human Factors in Electronics*, HFE-3, 43-51.

Costello, R. G. (1968). The surge model of the well-trained operator in simple manual control. *IEEE Transactions on Man-Machine Systems*, MMS-9, 2-9.

Crossman, E. R. and Cooke, F. W. (1974). Manual control of slow response systems. In E. Edwards and F. P. Lees (Eds.) *The Human Operator in Process Control*, London: Taylor and Francis, 51-66.

Cutting, J. E. (1986). *Perception with an eye for motion*. Cambridge, MA: The MIT Press.

Davies, R. (1957). The human operator as a single channel information system. *Quarterly Journal of Experimental Psychology*, 9, 119-129.

Gibson, J. J. (1947). *Motion picture testing and research*. (AAF Aviation Psychology Research Report No. 7) Washington, DC: U.S. Government Printing Office.

Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.

Gibson, J. J. (1962). Observations on active touch. *Psychological Review*, 69, 477-491.

Gibson, J. J. (1958). Visually controlled locomotion and visual orientation in animals. *British Journal of Psychology*, 49, 182-194.

Gibson, J. J., Olum, P., and Rosenblatt, F. (1955). Parallax and perspective during aircraft landings. *American Journal of Psychology*, 68, 372-385.

Goodman, N. (1968). *Languages of Art: An Approach to a Theory of Symbols*. Indianapolis: Bobbs-Merrill.

Hick, W. E. (1948). The discontinuous functioning of the human operator in pursuit tasks. *Quarterly Journal of Experimental Psychology*, 1, 36-51.

Jagacinski, R. J. (1977). A qualitative look at feedback control theory as a style of describing behavior. *Human Factors*, 19(4), 331-347.

Jagacinski, R. J. (In press). Target acquisition: Performance measures, process models, and design implications. In G. R. McMillan (Ed.) *Applications of Human Performance Models to Systems Design*. Plenum Press.

Jagacinski, R. J., Plamondon, B. D., and Miller, R. A. (1987). Describing movement control at two levels of abstraction. In P. A. Hancock (Ed.) Human Factors Psychology, New York: Elsevier.

Johnson, W. W., Bennett, C. T., O'Donnell, K., and Phatak, A. V. (1988). Optical variables useful in the active control of altitude. Paper presented at the *23rd Annual Conference on Manual Control*, Cambridge, MA, June.

Kleinman, D. L., Baron, S. , and Levison, W. H. (1970). An optimal control model of human response. Part 1: Theory and validation. *Automatica*, 6, 357-369.

Kleinman, D. L., Baron, S. , and Levison, W. H. (1971). A control theoretic approach to manned-vehicle systems analysis. *IEEE Transactions on Automatic Control*, AC-16, 6, 824-832.

Knoop, P. A. (1978). Survey of human operator modeling techniques for measurement applications. AFHRL-TR-78-35. Brooks Air Force Base: Air Force Systems Command.

Koenderink, J. J. (1989). Optic flow. *Vison Research*, 26, 161-180.

Larish, J. F. and Flach, J. M. (In press). Sources of optical information useful for the perception of speed of rectilinear self-motion. *Journal of Experimental Psychology: Human Perception and Performance*.

Longuet-Higgens, H. C. and Prazdny, K. (1980. The interpretation of a moving retinal image. *Proceedings of the Royal Society of London*, 208, 385-397.

McClelland, J. L. and Rumelhart, D. E. (1986). *Parallel Distributed Processing*. Cambridge, MA: MIT Press.

McRuer, D. T. Allen, R. W., Weir, D. H., and Klein, R. H. (1977). New results in driver steering control models. *Human Factors*, 19(4), 381-397.

McRuer, D. T. and Jex, H. R. (1967). A review of quasi-linear pilot models. *IEEE Transactions on Human Factors in Electronics*, HFE-8(3), 231-249.

McRuer, D. T. and Krendal, E. S. (1974). Mathematical models of human pilot behavior. AGARD-AG-188.

McRuer, D. T. and Weir, D. H. (1969). Theory of manual vehicular control. *Ergonomics*, 12, 599-633.

Owen, D. H. and Warren, R. (1987). Perception and control of self-motion implications for visual simulation of vehicular locomotion. In L. S. Mark, J. S. Warm, and R. L. Huston (Ed.) *Ergonomics and Human Factors: Recent Research*, New York: Springer-Verlag.

Owen, D. H., Warren, R., Jensen, R. S., Mangold, S. J., and Hettinger, L. J. (1981). Optical information for detecting loss in ones own forward speed. *Acta Psychologica*, 48, 203-213.

Pew, R. W. and Baron, S. (1978). The components of an information processing theory of skilled performance based on an optimal control perspective. In G. E. Stemach (Ed.) *Information Processing in Motor Control and Learning*, New York: Academic Press.

Powers, W. T. (1978). Quantitative analysis of purposive systems. Some spadework at the foundations of scientific psychology. *Psychological Review*, 85, 417-435.

Stappers, P. J. (1989). Forms can be recognized from dynamic occlusion alone. *Perceptual and Motor Skills*, 68, 243-251.

Stappers, P. J., Smets, G. J. F., and Overbeeke, C. J. (1989). Gaps in the optic array: Toward richer operationalizations of optic array and optic flow. Paper presented at the *Fifth International Conference on Event Perception and Action*, Miami University, Oxford, OH, July 24-28.

Stoffregen, T. A. (1985). Flow structure versus retinal location in the optical control of stance. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 554-565.

von Foerster, H. (1984). Observing systems Seaside, CA: Intersystems Publications.

Warren, R. (1976). The perception of egomotion. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 448-456.

Warren, R. and McMillan, G. R. (1984). Altitude control using action-demanding interactive displays: Toward an active psychophysics. *Proceedings of the 1984 IMAGE III Conference* (pp. 405-415). Sponsored by the Air Force Human Resources Laboratory, Phoenix, AZ.

Warren, W. H. Jr., Morris, M. W., and Kalish, M. (1988). Perception of translational headings from optic flow. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 646-660.

Welford, A. T. (1952). The 'psychological refractory period' and the timing of high-speed performance, a review and a theory. *British Journal of Psychology*, 43, 2-19.

Wickens, C. D. (1984). *Engineering Psychology and Human Performance*. Columbus, OH: Merrill Publishing Co.

Young, L. R. (1969). On adaptive manual control. *IEEE Transactions on Man-Machine Systems*, MMS-10, 292-331.

Young, L. R. and Meiry, J. L. (1965). Bang-bang aspects of manual control in high-order control systems. *IEEE Transactions on Automatic Control*, AC-10, 336-341.
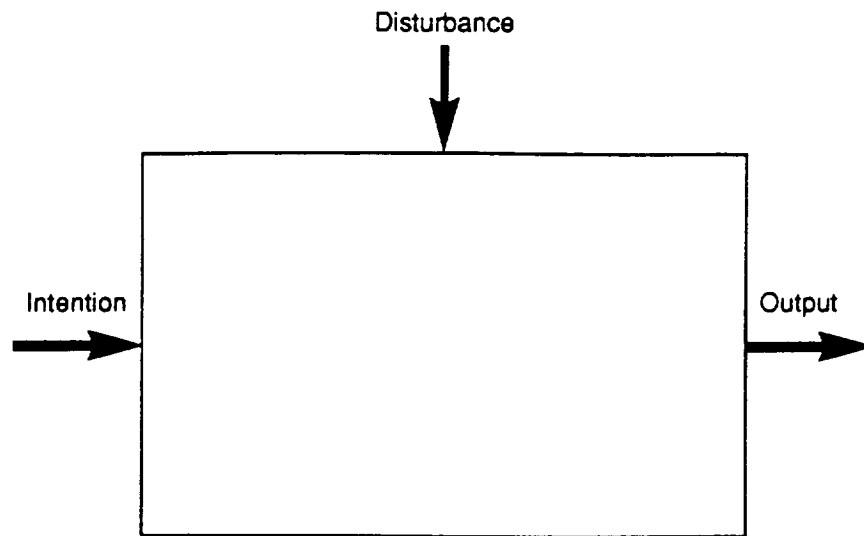
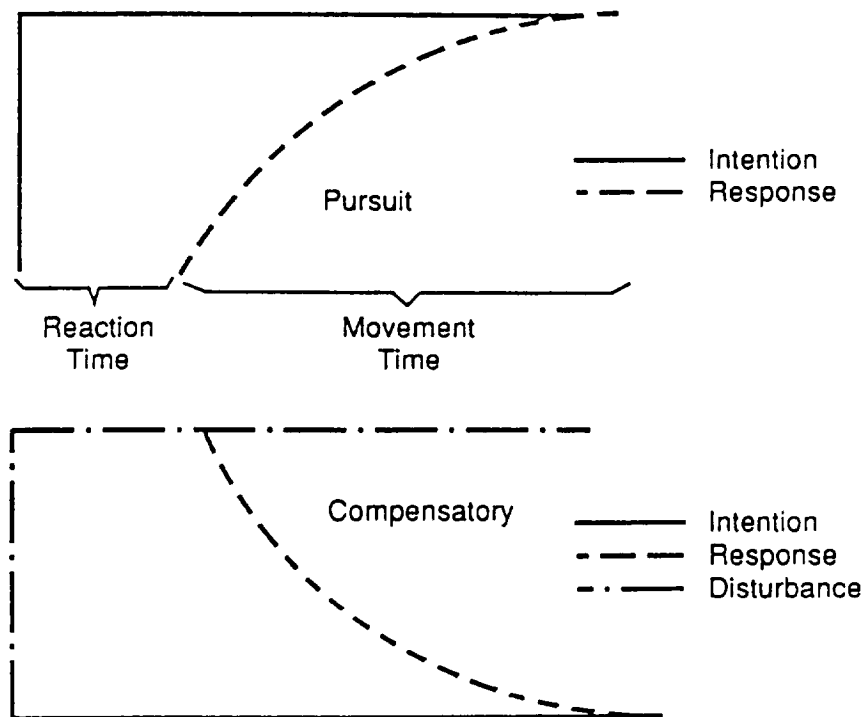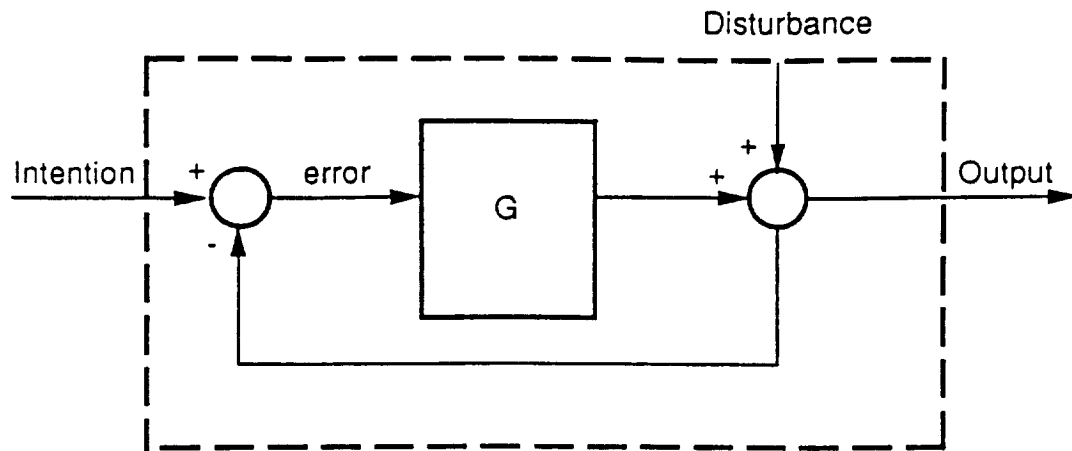Figure 1.   A black box representation of a human-environment system.

Figure 2.   Responses to a step input on the intention channel (pursuit) and on the disturbance channel (compensatory).
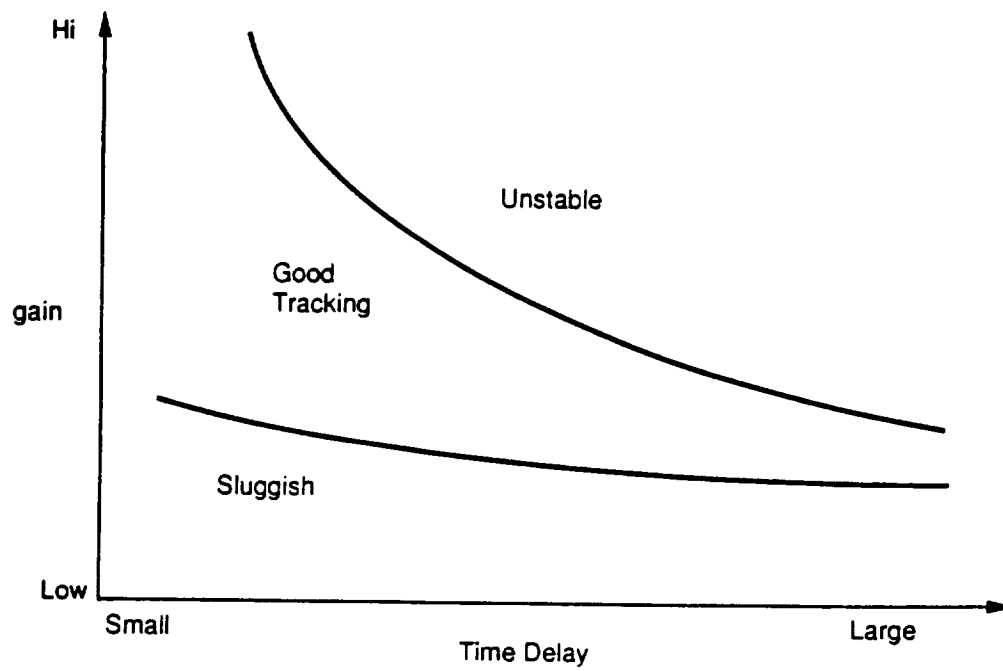
139

error = Intention - Output

G * error + Disturbance = Output

G (I - O) + D = O

GI - GO + D = O

GI + D = O + GO

GI + D = O (I + G)

$$\left[\frac{G}{1+G}\right]\text{Intention} + \frac{1}{1+G}\ \text{Disturbance} = \text{Output}$$

Figure 3. A simple negative feedback system.
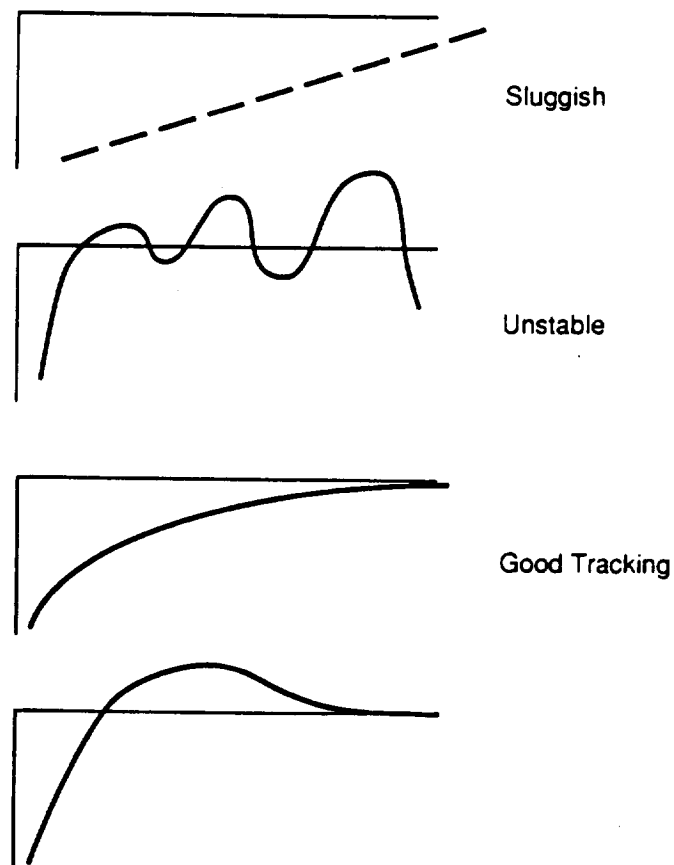
Figure 4. (A) Illustrates tracking quality as a function of gain (sensitivity to error) and the time delay (delay of feedback) (Adapted from Jagacinski, 1977) (B) Illustrates responses to step inputs for the regions shown in A.

141
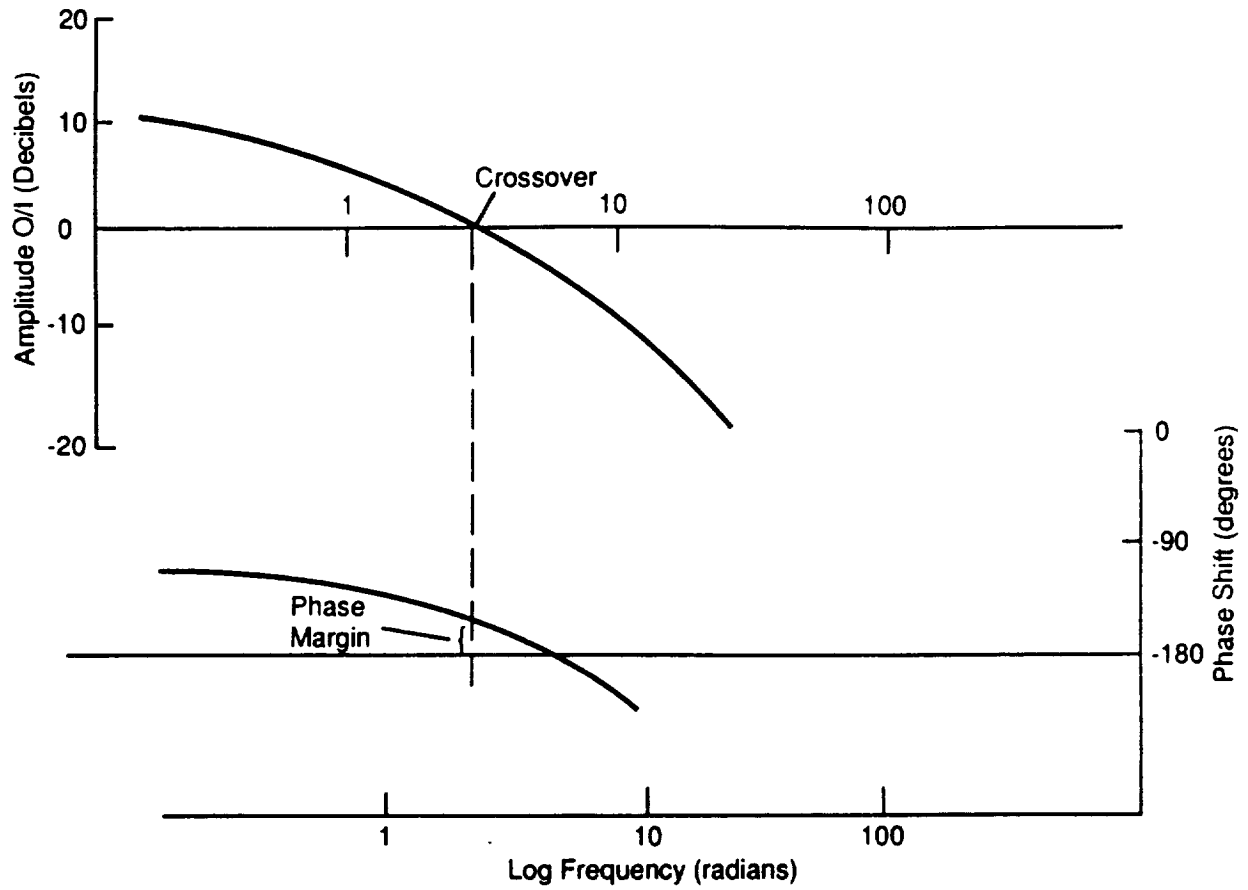
Figure 5. A Bode plot typical of a "good" controller.
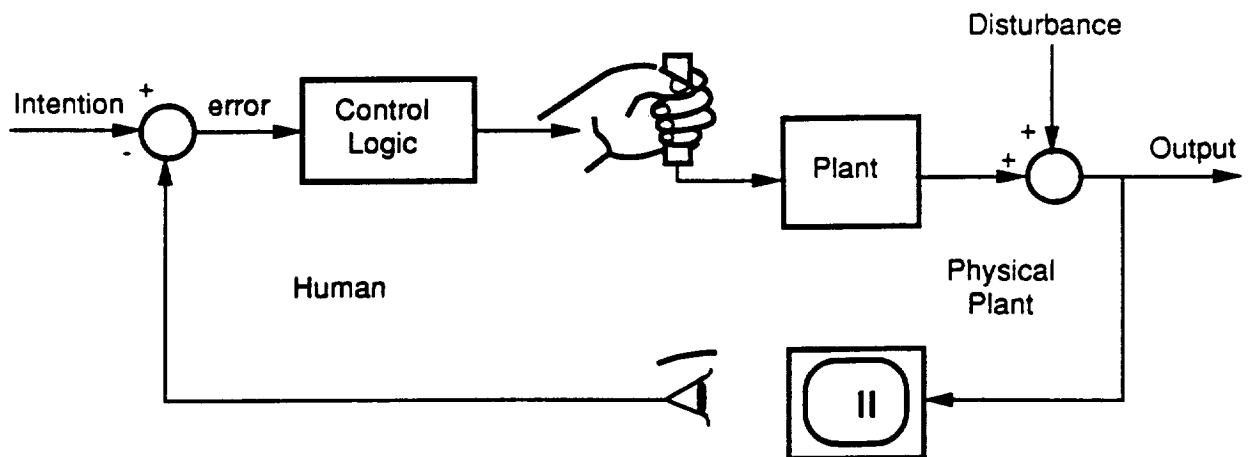
Figure 6. The manual control framework.
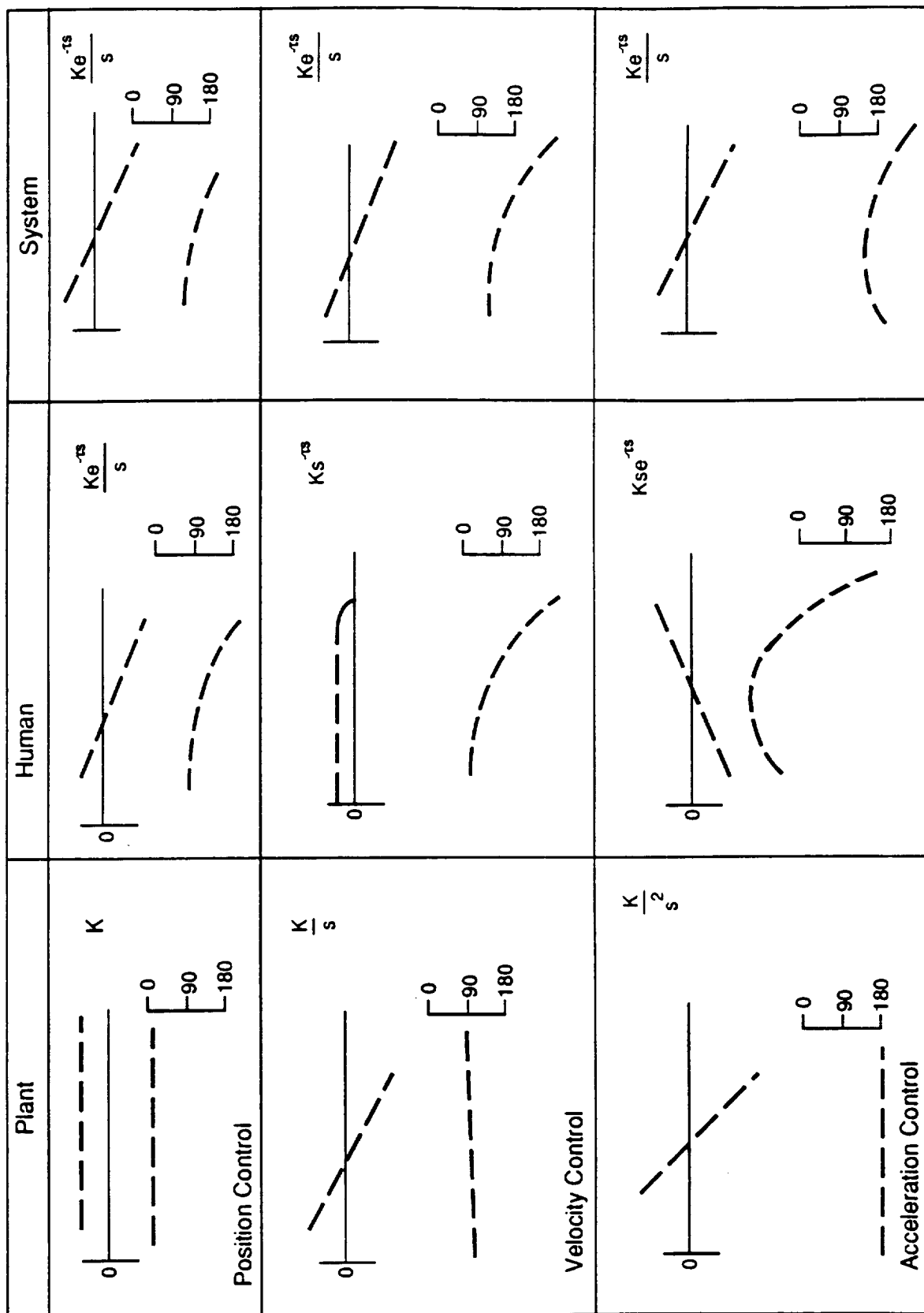
Figure 7. An illustration of the logic behind the crossover model. Shows how the human adapts to the demands imposed by three simple plants.

143
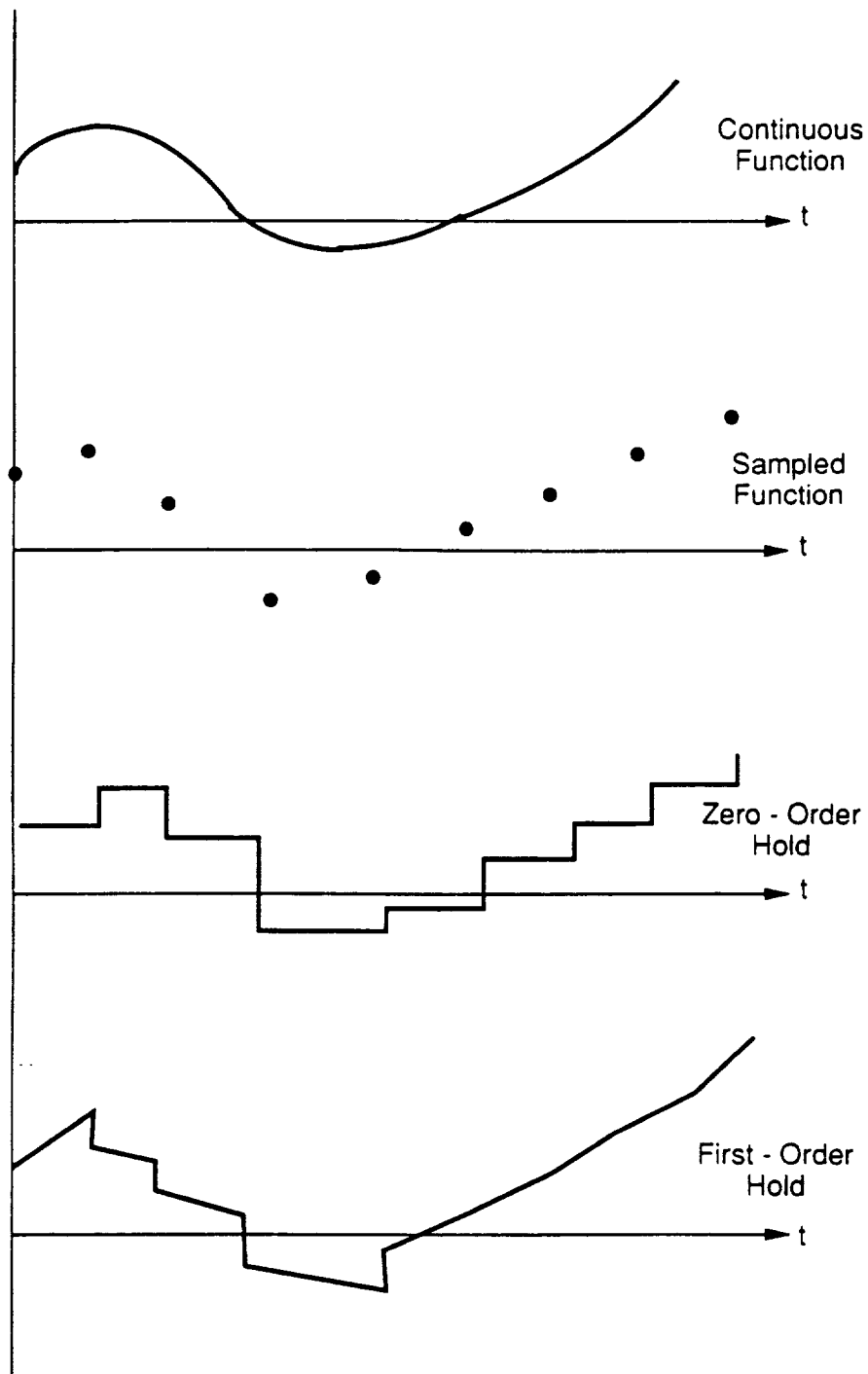
Figure 8. Two strategies for discrete synchronous control. Zero-order extrapolates based on position. First-order extrapolates based on position and velocity (Adapted from Bekey, 1962).

(a)

(b)
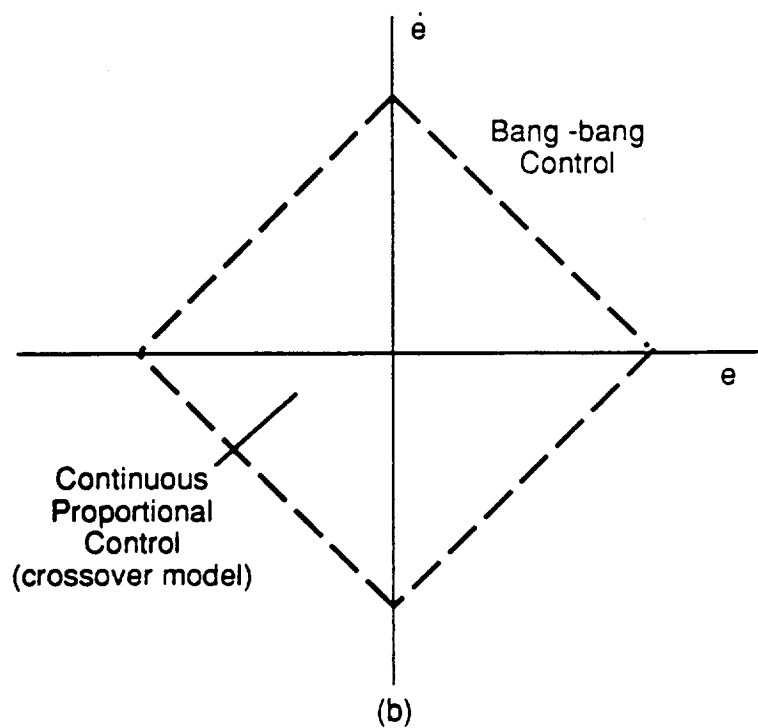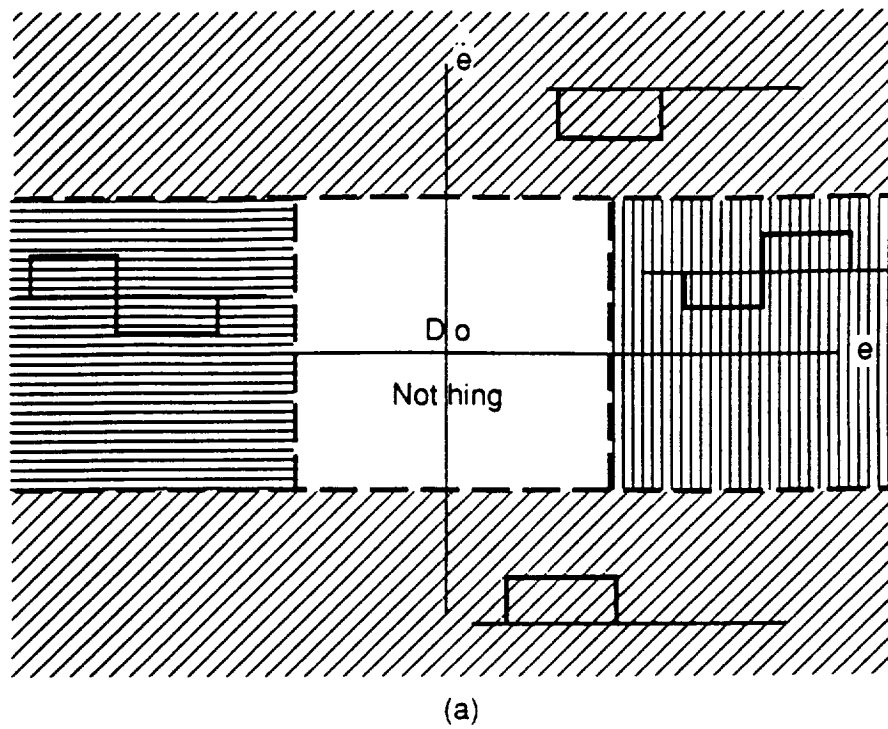
Figure 9. (A) Logic for an asynchronous discrete controller proposed by Angel and Bekey (1968) (B) Logic for hierarchical "surge" model proposed by Costello (1968).
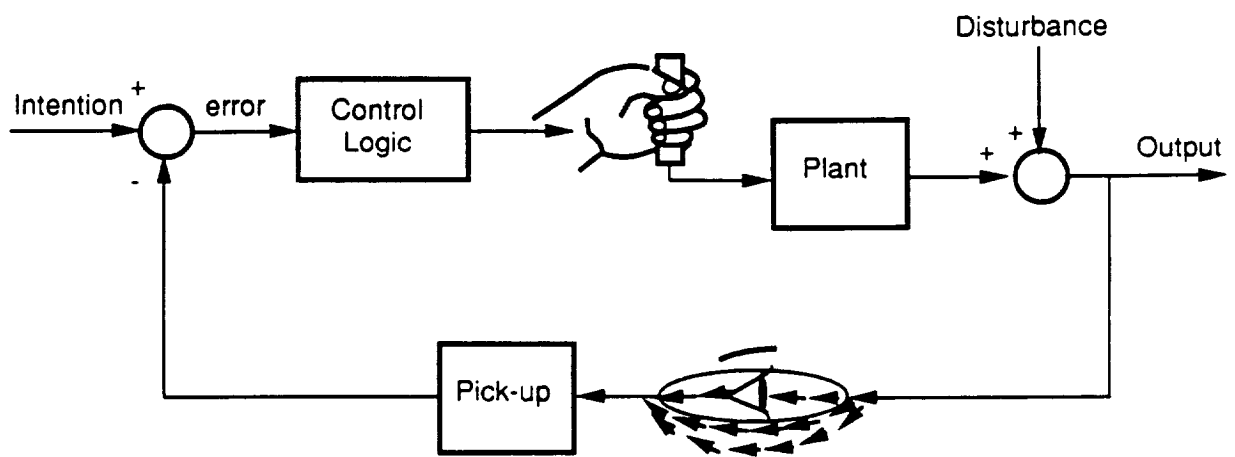
Figure 10. Closing the loop through the optic array.
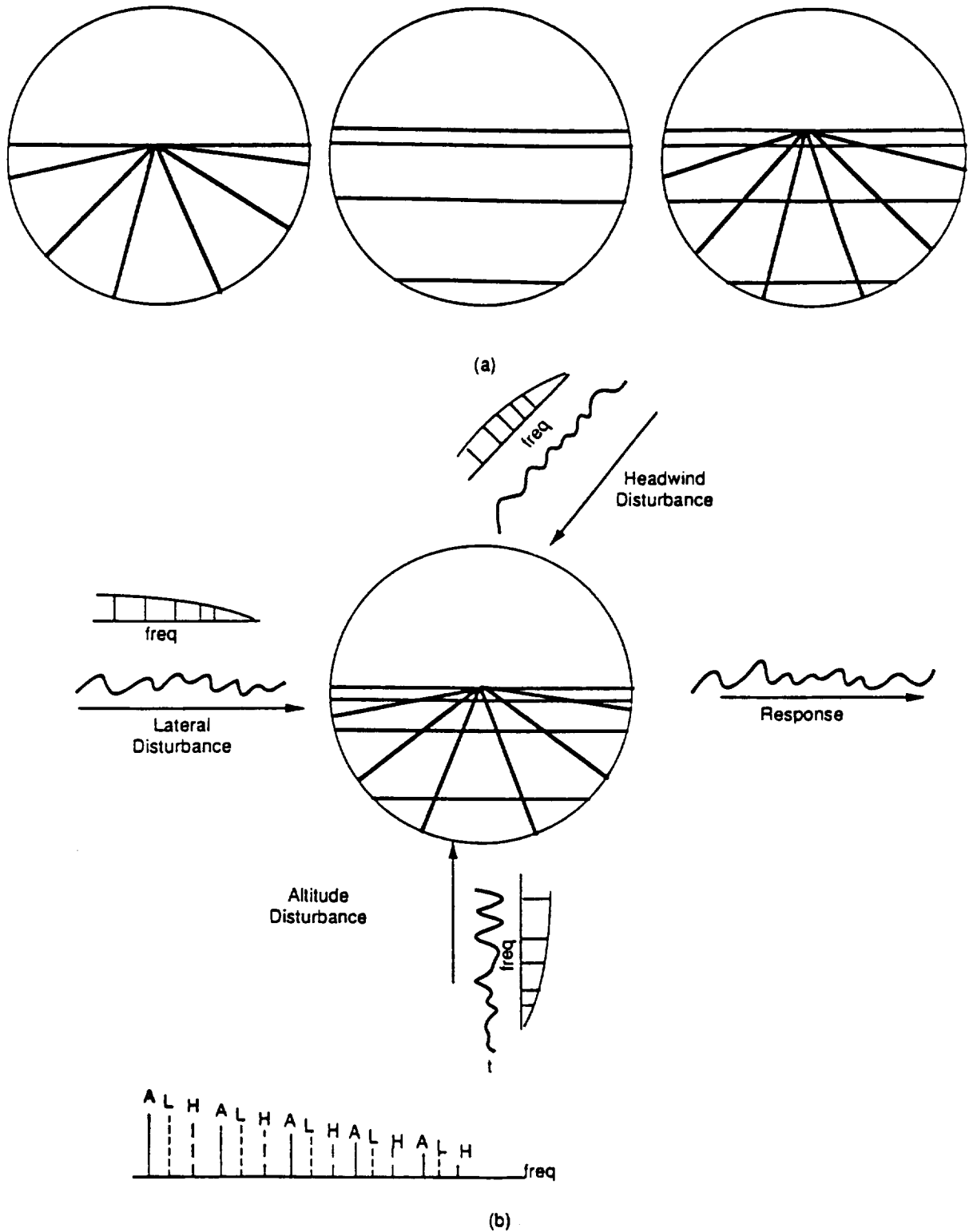
Figure 11. Illustrates logic of approach employed by Johnson et al. (1988) to evaluate alternative invariants for altitude control (A) Parallel (splay) texture, perpendicular (density) texture, and square texture (B) Frequency is used as a signature to isolate the effects of three disturbances (altitude, head wind, lateral) that were chosen because of their specific impacts on parallel and perpendicular texture.
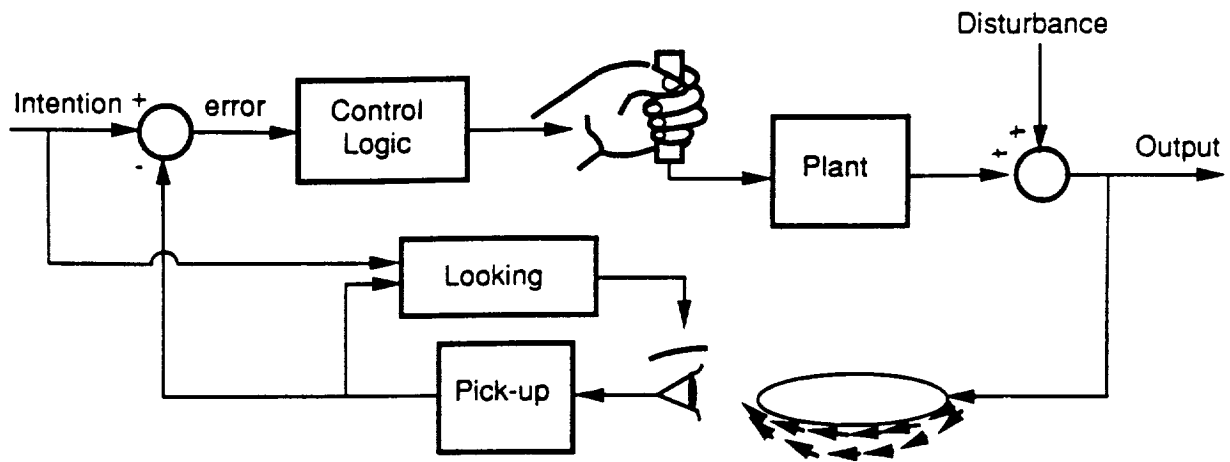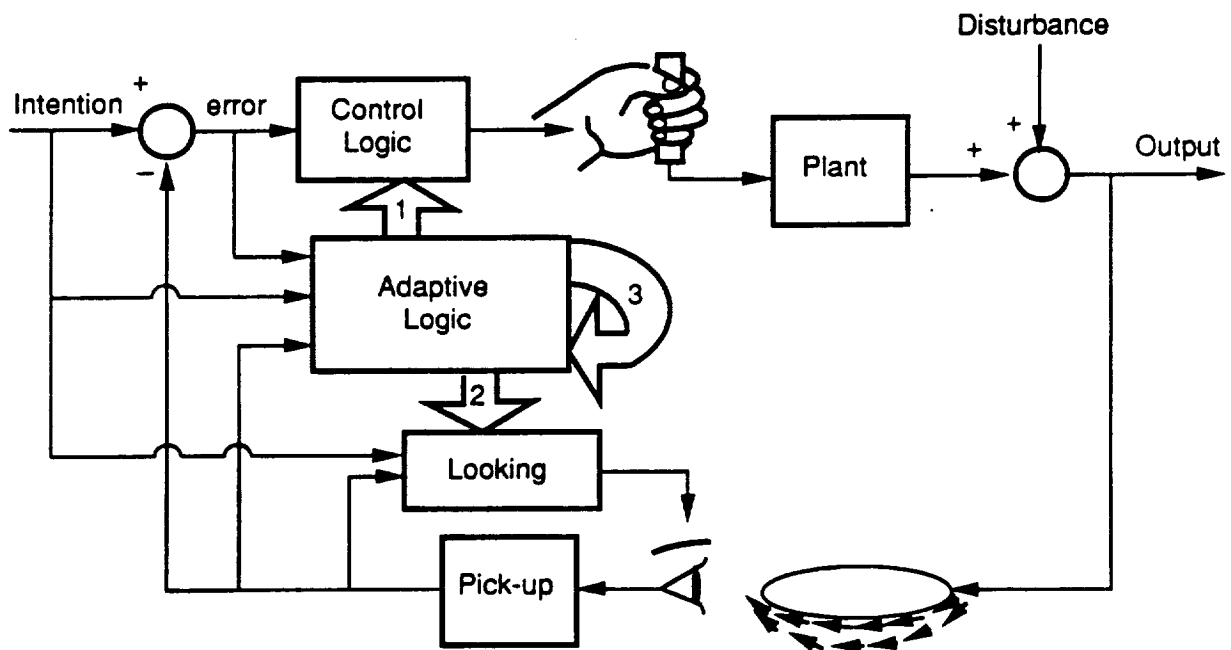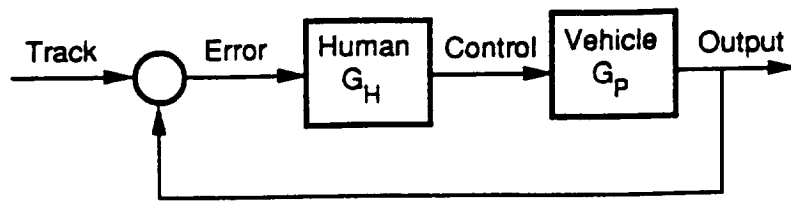
Figure 12. Uncoupling the eye (exploratory mode) from the hand (performatory mode).

1. Changing action strategies (motor learning)
2. Changing search strategies (discrimination learning)
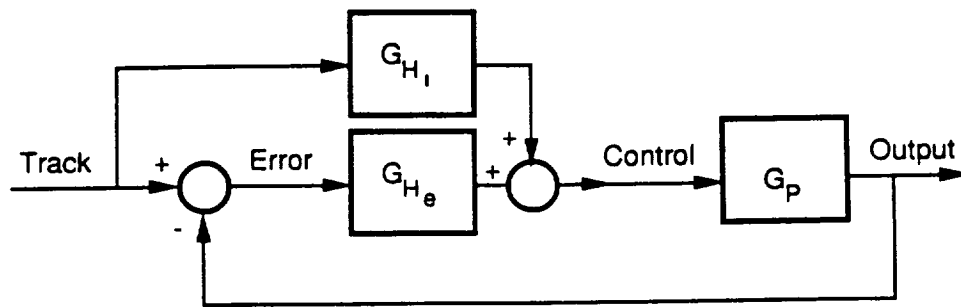3. Changing adaption strategies (learning to learn)

Figure 13. Adaptation—operating on operators.

(a)

Track → Error → Human G_H → Control → Vehicle G_P → Output

$$\left[\frac{G_H G_P}{1 + G_H G_P}\right] \text{Input} = \text{Output}$$

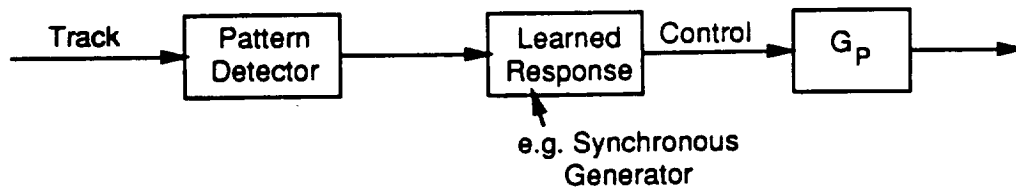Note: Match between input and output depends on open loop gain ($G_H G_P$)

(a)

(b)

G_{H_i}

Track → + Error → G_{H_e} → + + → Control → G_P → Output

$$\left[\frac{G_P (G_{H_i} + G_{H_e})}{1 + G_{H_e} G_P}\right] \text{Input} = \text{Output}$$

Note: If $G_{H_i} = \frac{1}{G_P}$ (perfect mental model of plant), then input will equal output regardless of the open loop gain of the compensatory loop ($G_{H_e} G_P$)

(b)

(c)

Track → Pattern Detector → Learned Response → Control → G_P →

e.g. Synchronous Generator

Note: Open - loop control

(c)

Figure 14. The Successive Order of Perception model (SOP) proposed by McRuer et al. (1977) includes three control modes (a) compensatory, (b) pursuit, and (c) precognitive.

149

CONTEXTUAL SPECIFICITY IN PERCEPTION AND ACTION

Dennis R. Proffitt
University of Virginia
Charlottesville, Virginia

The visually guided control of helicopter flight is a human achievement, and thus,understanding this skill is, in part, a psychological problem. The abilities of skilled pilots are impressive, and yet it is of concern that pilots' performance is less than ideal:They suffer from workload constraints, make occasional make errors, and are subject to such debilities as simulator sickness. Remedying such deficiencies is both an engineering and a psychological problem.

When studying the psychological aspects of this problem, it is desirable to simplify the problem as much as possible, and thereby sidestep as many intractable psychological issues as possible. Simply stated, we do not want to have to resolve such polemics as the mind-body problem in order to contribute to the design of more effective helicopter systems. On the other hand, the study of human behavior is a psychological endeavor and certain problems cannot be evaded.

In this paper I discusses four related issues that are of psychological significance in understanding the visually guided control of helicopter flight. First, I present a selected discussion of the nature of descriptive levels in analyzing human perception and performance. Here I will argue that the appropriate level of description for perception is kinematical, and for performance, it is procedural. Second, I argue that investigations into pilot performance cannot ignore the nature of pilots' phenomenal experience. The conscious control of actions is not based upon environmental states of affairs, nor upon the optical information that specifies them. Actions are coupled to perceptions. Third,I discuss the acquisition of skilled actions in the context of inherent misperceptions. Such skills may be error prone in some situations, but not in others. Finally, I discuss the contextual relativity of human errors.

Each of these four issues relates to a common theme: The control of action is mediated by phenomenal experience, the veracity of which is context specific.

LEVELS OF DESCRIPTION

How do we characterize what helicopter pilots are doing? One answer to this question is that pilots are controlling the dynamics of their craft. At one level of description it makes sense to describe pilots' behavior in these terms; however, at another it does not.

Within some control theory models, pilots are characterized as having perfect understandings of the dynamical properties of their flight environment. Given the nature of the variables under examination in these models, it does not matter that the achievement of such dynamical understandings

makes no psychological sense. However, from the point of view of understanding task performance related to other pilot variables, more appropriate characterizations of human behavior are needed.

I argue that pilots can achieve only very simplistic understandings about dynamics, and that the appropriate level of description for perception is kinematical, and for control, it is procedural.

Dynamics Versus Kinematics

In Classical Mechanics, dynamical analyses relate to the action of bodies that move due to the application of forces. In nature, object motions are constrained by the law of least action, where action has the dimensions of energy x time. Kinematical analyses, on the other hand, deal only with pure object motions without consideration of mass, and thus, of energy. As a level of analyses, kinematics is far less restrictive than is dynamics: Most of the object motions that can be describe in kinematics are inconsistent with Newton's Laws.

Research has shown that people's understandings of dynamics is extremely simplistic and heuristical (Proffitt & Gilden, 1989). Moreover, the spontaneous dynamical intuitions of trained physicists differ very little from those of unsophisticated people. Physicists' expertise becomes evident only when they are permitted to symbolically represent the system under consideration. In this sense physicists have a dual awareness: One is immediate, appeals to phenomenal categories, and differs little from naive common sense; the other is deliberate, appeals to the symbolic categories of first principle representations (e.g. $F = ma$), and is far removed from common sense. I propose that helicopter pilots do not fly their crafts by controlling dynamics. Being people, pilots have neither the perceptual nor conceptual ability to penetrate their helicopter's dynamics during flight. Rather the problem of representing the control of helicopter flight is best stated in terms of a mapping between phenomenal variables. That is, pilots must relate the kinematical variables available in perceptual stimulation to appropriate control actions. The dynamics of the craft constrain the nature of this perception/action coupling; however, the pilot need not appreciate these dynamics in order to exploit them. In essence, pilots need to appreciate the dynamics of helicopters no better than children need to understand the dynamics of their bicycles. The rules that define skilled control of a particular mechanical system, need not embody any of the system's dynamics. These rules (transform functions) relate one class of kinematical variables, perceptions, to another, actions.

Declarative Versus Procedural Knowledge

There is a very old distinction between "knowing how" versus "knowing that" that has more recently come to be described as procedural versus declarative knowledge. Procedural knowledge consists of rules for regulating skilled behaviors; they are recipes for action, are evoked by specific situational variables, and are typically not accessible to awareness. Riding a bicycle or flying a helicopter depends upon procedural forms of knowledge. Declarative knowledge is explicit and entails a conscious conceptualization and articulation about some state of affairs.

Piloting a helicopter evokes procedural knowledge. These rules are not general because they are blind to the underlying dynamics of the vehicle. The dynamics of helicopter flight create an environment in which particular kinematical variables in perception and action are related in specific

ways. Learning these relationships establishes procedures for producing desired kinematical outcomes.

PHENOMENA

Pilots fly helicopters by heeding and affecting phenomenal states of affairs. What are the relevant phenomena? During the NASA Workshop, my group picked slant perception as a phenomena to study.

This choice was motivated by the existence of a striking everyday phenomena that may jeopardize successful low altitude flying. When, for example, people drive in San Francisco, they cannot help but be struck by the incredibly steep inclines of some of the roads that they encounter. When asked to estimate the slopes of these hills,people provide erroneous estimates in the neighborhood of 45-75 deg (informally collected anecdotal evidence). In fact, the steepest road is no more than 15 deg. Evidence exists that this is a general finding (Ross, 1974). When approaching a large incline, such as a hill, people grossly overestimate its slant.

We decided to study the psychophysics of this phenomena by initially asking the question: What slant will be perceived for (1) various hill slants, (2) viewed at various altitudes, (3) by a moving observer who either approaches or moves laterally with respect to the hill (4) at different speeds. These, of course, are frequently encountered situations for helicopter pilots.

Our prediction was that slant will be greatly overestimated in all conditions and that this error will be greatest when the hill is approached head-on at low altitudes. Other more specific hypotheses were formulated for each of the other variables.

In addition to mapping out the psychophysics of slant perception across these variable, we hope to determine the visual variables that affect slant perception, and ultimately to develop a model for human slant perception. With regard to this latter goal, levels of description issues again emerge. In particular, we would like to know the geometrical space in which kinematic information is represented (see Lappin this volume).

From a geometrical perspective, the slant of a hill is fully specified to a moving observer; however, people seem not to appreciate well the optical information that is available. This implies that people have either (1) little sensitivity to the available information, or (2) that they possess the required sensitivities, but are unable to use it effectively when making slant judgments.

CONTROL

Given that people misperceive the slant of hills, why do they not evidence this misperception when walking up them? The answer to this question, and the related helicopter control issue, is that accurate perceptions of environmental state of affairs are not required for effective control of action

within the situation. So long as perceptual attributes co-vary perfectly with environmental dimensions, control will not reflect on underlying misperceptions.

Consider how control behaviors are learned in a situation like flying a helicopter at low altitudes over a hill. Suppose that the novice pilot misperceives the slant of a hill to be 60 deg when, in fact, its inclination is 15 deg. In order to maintain the desired altitude relative to the hill, the pilot must learn to couple the appropriate control responses to what is perceived. To put the matter simply, he or she must learn to pull back on the stick by some amount, given that a hill of some perceived slant is approaching. Through learning, the pilot will come to couple the appropriate control responses to the misperception of slant. It does not matter that slant is misperceived, since control responses have been acquired in the context of this misperception, and the misperception co-varies with distal slant.

That fundamental misperceptions may not be evidenced in particular control contexts, does not imply that they will never result in pilot error. One working hypothesis for the overestimation of slant is a conjecture that the perceived horizon is displaced below its actual location. This concomitant to slant misperception might have no influence on flying over a hill, but might very well effect judgments of the height for obstacles encounter on the hill. Given that the perceived location of the horizon may serve as an important cue to whether an obstacle is above or below ones flight path, misperceiving slant may result in errors in some contexts but not in others.

ERRORS

The sorts of control errors that people make tend to be context specific. Accidents that occur in helicopter flight are known to be far more likely in certain situations than in others. Assuming a skilled operator, the contextual specificity of control errors derives not from a single cause but rather from at least three quite different sources.

Workload

Obviously, some situations require considerably more effort than do others. Some of these contexts present a greater diversity of task relevant information requiring attentional allocation and information integration. In other situations, the control behaviors are particularly arduous. And finally, some situations present especially difficult demands on both perceptual and control resources.

Degraded or Missing Information

As tasks move farther from those encountered in everyday experience, it becomes increasing likely that the information that is typically relied upon to perceive some environmental state of affairs may be reduced or absent. For example, optical flow rate specifies speed only if the observer knows his or her altitude. Thus, optical flow suffices in perceiving speed for a locomoting person accustomed to his or her own eye height, but not when that person is piloting an aircraft.

154

Misperceptions

As the above discussion on slant misperception noted, misperceptions may be inconsequential in some contexts, but not in others. Moreover, context can be defined in two quite different ways. First, context may be defined in terms of the environment: maintaining a constant altitude while flying over a hill versus deciding whether a tree is above or below one's flight path. Here the contexts have an external referent: hills and trees. On the other hand, contexts can be defined by differential task demands that arise in the same physical situation. Thus, for example, a pilot may successfully maintain a constant altitude while piloting his or her craft over a hill, thereby implying that the hill's slope was accurately perceived. However, if asked to estimate the slope of the hill verbally, or by adjusting a visual or manual slant indicator, that same individual will evidence a strong overestimation of slant.

It is tempting to ignore or disparage the significance of the explicit slant estimation error, since only the control of altitude has practical significance. I think that this would be a mistake. If we want to understand what a pilot is doing, we must take a psychological perspective, and thereby recognize that the visually guided control of action is mediated by phenomenal experience. Thus, an adequate account of visually guided control cannot simply attempt a mapping of environmental properties, as they are manifest in optical structure, onto control behaviors. Visual experience is formed by optical structure, but it is not equivalent to it. To assume otherwise is a futile attempt to sidestep the difficult issues inherent to the study of human behavior.

CONCLUSION

The dynamics of helicopter flight create an environment. In this environment, pilots learn to relate particular kinematical variables in perception and action. Learning consists of discovering how control procedures transform current phenomenal states into those desired in the future. These procedures cannot be general, since they were acquired without any first-principle understanding of the dynamics inherent in the context of their acquisition. In addition, control procedures are often acquired in the context of misperceptions. Yet, because of their contextual specificity, they may lead to errors in only a limited set of situations.

REFERENCES

Proffitt, D.R. & GIlden, D.L. (1989). Understanding natural dynamics. Journal of Experimental Psychology: Human Perception and Performance, 15, 384-393.

Ross, H. (1974). Behavior and perception in strange environments. London: Allen and Unwin.

VISUALLY GUIDED CONTROL OF MOVEMENT IN THE CONTEXT OF MULTIMODAL STIMULATION

Gary E. Riccio
University of Illinois
Urbana-Champaign, Illinois

ABSTRACT

Flight simulation has been almost exclusively concerned with simulating the motions of the aircraft. Physically distinct subsystems are often combined to simulate the varieties of aircraft motion. "Visual display systems" simulate the motion of the aircraft relative to remote objects and surfaces (e.g., other aircraft and the terrain). "Motion platform" simulators recreate aircraft motion relative to the gravitoinertial vector (i.e., correlated rotation and tilt as opposed to the "coordinated turn" in flight). "Control loaders" attempt to simulate the resistance of the aerodynamic medium to aircraft motion. However, there are few operational systems that attempt to simulate the motion of the pilot relative to the aircraft and the gravitoinertial vector. The design and use of all simulators is limited by poor understanding of postural control in the aircraft and its effect on the perception and control of flight. Analysis of the perception and control of flight (real or simulated) must consider that (a) the pilot is not rigidly attached to the aircraft and (b) the pilot actively monitors and adjusts body orientation and configuration in the aircraft. It is argued that this more complete approach to flight simulation requires that multimodal perception be considered as the rule rather that the exception. Moreover, the necessity of multimodal perception is revealed by emphasizing the complementarity rather than the redundancy among perceptual systems. Finally, an outline is presented for an experiment to be conducted at the NASA Ames Research Center. The experiment explicitly considers possible consequences of coordination between postural and vehicular control.

1.0 AN EXOLOGICAL PERSPECTIVE ON FLIGHT SIMULATION

1.1 Purpose and Assumptions

One purpose of research in flight simulation is to enhance the simulation of the force and motion environment generated by an aircraft. A need for enhancements is based largely on the assumption that extant systems do not adequately simulate certain flight regimes. The criteria for adequacy are rarely stated explicitly. The implicit criteria fall into two general categories: (a) Subjective experience in the simulator and the aircraft should be similar. Ideally, the simulation should not be perceived as such, but rather as motion of the pilot in an environment with recognizable objects. (b) Flight control skills acquired in the simulator and those acquired in the aircraft should be similar. Ideally, transfer of training from the simulator to the aircraft should be cost effective.

Inadequacy is an assumption because there has not been sufficient formal experimentation to conclude that any flight simulator is inadequate. However, it is equally important that there has not been sufficient formal experimentation to conclude that any flight simulator is adequate (cf. Cardullo & Sinacori, 1988; Lintern, 1987). The dearth of formal experimentation on the adequacy of flight simulators is almost certainly due to the fact that the criteria for adequacy are considered to be too nebulous or too complex in any situation that even remotely resembles flying an aircraft. Because of this fundamental lack of information, there has been considerable speculation and controversy about the utility of various flight simulation systems. In spite of the lack of information, there have been developments in flight simulation. One of the challenges for research in flight simulation is to demonstrate that new simulation concepts can be derived within a substantial scientific framework.

1.2 Approach

Developments in flight simulation have relied primarily on "sound engineering judgment," that is, on the ability of the engineer to translate the needs of the user into the actions of some physically realizable system. While this process can be very efficient, its effectiveness is limited by the precision (detail) and accuracy (validity or relevance) of specifications provided by the user. Developments in flight simulation may not engender improvements in usefulness if they are motivated by specifications that are not relevant to explicit criteria for adequacy. This is especially problematic in the design of human-machine systems because of the limited capacity for analytic introspection (by novices or experts) about the factors that are relevant to perception and action.

A more tractable approach to flight simulation has been to focus on the "limiting factors" in flight control that are peculiar to the simulator. The focus in on the interactions between the perception and control of the aircraft's attitude and motion, that is, the way in which perception of the aircraft's attitude and motion influences control of the aircraft's attitude and motion. Other factors (e.g., orders, plans, and threats) influence the pilot's actions once the situation is perceived, but such factors are more or less arbitrary given the plethora of present and future flight scenarios. Moreover, such factors must take into account the constraints on observability and controllability imposed by the human-machine system. This has provided a "principled basis" for developments in flight simulation: developments should be motivated by theory and experiments in psychophysics and manual control that suggest the ways in which observability and controllability of attitude and motion is different in the simulator and the aircraft. This approach is exemplified by ecological and control-theoretic research in flight simulation (e.g., Kron, Cardullo, & Young, 1980; Flach, Riccio, McMillan, & Warren, 1986; Martin, McMillan, Warren, & Riccio, 1986; Cardullo & Sinacori, 1989; Warren & Riccio, 1986; Riccio & Cress, 1986; Riccio, Cress, & Johnson, 1987; Riccio 1989; Riccio & Stoffregen, 1988, 1989; Stoffregen & Riccio, 1988, 1989a; Zacharias, Warren & Riccio, 1986).

It is sometimes suggested that developments in flight simulation could be based on "cognitive theory" or "consistent pilot opinion," but no principled basis for inclusion of such factors has ever been revealed. Cognitive theory should be dismissed as a basis for developments in flight simulation because it reveals virtually nothing about limitations that are peculiar to the simulator. One could consider situation-specific anxiety (e.g., about crashing) that may not be present in extant simulators; however, anxiety inducing devices in flight simulators have never been considered seriously. Any other differences between cognition in the simulator and in the aircraft are ultimately attributable to differences in observability and controllability. Pilot opinion is also questionable as a basis for

developments in flight simulation. It should not be considered seriously unless there is corroborating theory suggests an important role for a particular source of information but where experimental evidence is either unavailable or inconclusive.

1.3 Unique Areas of Emphasis

The *sine qua non* of flight simulation is generally considered to be the capacity to induce perception of self motion through an environment without moving the observer. This capacity becomes useful if the observer is allowed to control the simulated self motion; that is, the observer-actor can achieve goals. Most goal directed motion through the environment requires perception of objects and surfaces that are distant from the observer. Visual perception is thus crucial for goal directed motion. For this reason, there is no question that "visual display systems" are necessary in flight simulation. There is general agreement that further developments in visual display systems are important because recognition of familiar objects and layouts increases the range of flight tasks that can be performed in the simulator. For example, the detail on a tanker aircraft is important in the approach and docking phases of in-flight refueling; the depth of a ravine or the presence of telephone wires is important in low level flight. In addition, there is no question that visual display systems are sufficient to induce the perception of constant velocity or low acceleration. The issue in flight simulation over which there is the greatest controversy, and for which there is the greatest design consequences, is whether there are any situations where visual display systems are not sufficient (e.g., Cardullo & Sinacori, 1988; Lintern, 1987).

The design considerations in flight simulation can be organized into three categories: movement of the aircraft relative to an inertial reference frame (section 1.3.1), management of kinetic and potential energy (section 1.3.2), and coordination of postural and vehicular control (section 1.3.3). Modifications to extant flight simulators are suggested in each of these categories. The basis for the modifications is provided by a consideration of the exigencies for perception and control. The relevant interactions between perception and control are summarized in conceptual block diagrams (see Fig. 1, Fig. 2, and "Glossary").

1.3.1 Movement relative to an inertial reference frame. The focus here is on acceleration. Motion cannot be controlled without producing variations in velocity. Goal directed motion requires that these variations are observable. The question for flight simulation is whether these variations (i.e., acceleration) can be perceived visually, and if so, whether these variations (i.e., acceleration) can be perceived visually, and if so, whether they are attributed to motion of the environment or motion of the observer. It is important to note that there is very little research that is relevant to this issue. The basic research on visual perception of acceleration generally concentrates on object motion. Basic research on the visual perception of egomotion generally involves situations where acceleration if either small, nonexistent, or irrelevant to the task. Moreover, the visual perception of accelerative self motion is rarely mentioned as a theoretically important issue. It is especially surprising that the visual perception of vehicular acceleration has been largely neglected in flight simulation research.

If the visual perception of vehicular acceleration were in some way deficient, it would be important to exploit vestibular and somatosensory perception in flight simulation. The sensitivity of these systems to acceleration is well established. In this respect it is important to note that deficiencies in

159

the visual perception of vehicular acceleration would not necessarily be due to limitations in the visual system. Such deficiencies may exist because vehicular acceleration is fundamentally a multimodal phenomenon. By was of analogy, perception of vehicular acceleration without multimodal stimulation (i.e., with only the visual system) may be like perception of color without stimulating the "cone" cells of the retina (i.e., with only the "rod" cells). The visual perception of accelerative self motion may be limited (like the function of rod cells) to low levels of stimulation, perhaps as in special cases of postural sway (Stoffregen & Riccio, 1989b).

The most obvious concern about excessive reliance on visually simulated self motion is that the phenomenon requires the presence of optical structure. Optical structure is not always available in flight (e.g., at night, under a uniform sky, over water). Use of simulators is potentially more important in these dangerous conditions than in good visual conditions. Nonvisual stimulation would not be an option, it would be a necessity, if the simulator were to be used in such optically impoverished situations. A challenge for developments in flight simulation is to design systems that provide information about vehicular acceleration without relying on the visual system.

1.3.2 Management of kinetic and potential energy. The focus here is on coordinated maneuvers. An approach that is based on coordinated maneuvers is to be contrasted with one that is based on the degrees of freedom that can potentially be controlled independently in an aircraft. For example, the so-called "degree-of-freedom" approach might consider perception of roll, pitch, yaw, and airspeed to be fundamental (lift, drag, and thrust might be considered most fundamental but they would be difficult to relate to perceptual sensitivity). Data on the sensitivity of perceptual systems to these degrees of freedom of motion could be exploited in the design and integration of visual and nonvisual "display" systems for flight simulators. The advantage of the degree-of-freedom approach is that there is a considerable body of basic research that can be used to quantify the design process and objectify design decisions. However, there are several disadvantages to this approach: (a) an additional step is needed to reduce these data to a form that directly relates to actual flight control tasks (i.e., maneuvers); (b) there may be interactions among the degrees of freedom that alter sensitivity to the individual degrees of freedom of motion; (c) new dimensions of control may emerge when motions in various degrees of freedom covary.

A "maneuver based" approach would consider the aircraft's trajectory or flight path through the environment to be more basic than the mediate control parameters. Control of the trajectory involves changes in altitude and heading that constrain the covariation among roll, pitch, yaw, and airspeed. (It follows that adjustments of the stick, rudders, and throttle are also constrained to particular patterns of covariation.) The way in which covariation is constrained depends on the "evaluation function" for control. While the function (or criteria) on which control is evaluated (or guided) can vary, a generally important criterion that guides control is energy management. With respect to this criterion, efficient flight requires that the pilot monitor (directly or indirectly) the kinetic and potential energy of the aircraft. In particular, the pilot should be sensitive to the rate of change in, and exchange between, these parameters.

Management of kinetic energy requires control of the aircraft's velocity. The issues that pertain to perception of changes in velocity were mentioned above. Management of potential energy involves control of the so-called "G" forces acting on the aircraft. The magnitude and direction of the G forces are controlled primarily in curved trajectories (e.g., a "pull up" or a "coordinated turn").

The curvature of the trajectory determines the magnitude of the G forces. The attitude with respect to the trajectory (e.g., "angles of attack") determines the direction of the G forces on the aircraft. The magnitude and direction of the G forces, in turn, influences the trajectory of the aircraft. It is not known to what extent perceiving the magnitude and direction of G forces is required to produce efficient (coordinated) trajectories. Since the G forces are lawfully related to the radius and orientation of the trajectory, perceiving the trajectory kinematics could be sufficient. In principle, kinematic information is available to the visual system whenever optical structure is available. The question for flight simulation is whether the radius and orientation of the aircraft trajectory can be perceived visually. Again, the paucity of relevant data is noteworthy. This is surprising since the relevance of trajectory radius extends beyond flight control (e.g., perception of trajectory radius for the head would be useful in understanding the coordination of body segments during stance and pedal locomotion; Riccio & Stoffregen, 1988).

If the visual perception of trajectory radius and orientation were in some way deficient it would be important to exploit vestibular and somatosensory perception in flight simulation. The relationship between canal and otolith stimulation would seem ideally suited for perception of trajectory radius (unfortunately there are few data that directly relate to this hypothesis; Riccio & Stoffregen, 1989). There would be important implications for simulator design if people were actually sensitive to this relationship. perception of G forces could substitute for perception of trajectory radius and orientation. The sensitivity of vestibular and somatosensory systems to the direction and magnitude of G forces is not controversial (although the basis for this sensitivity is in question; Howard, 1986; Stoffregen & Riccio, 1988, 1989a; Riccio 1989).

It should be noted that curved trajectories are fundamentally multimodal phenomena. Again, an analogy to color vision may be useful. Instead of the electromagnetic "spectrum," the relevant continuum would be trajectory radius. Pure linear motion would be at one end of the continuum and pure angular motion at the other. Different kinds of sensors (i.e., with ranges of sensitivity to motion that differ with respect to their dependence on trajectory radius) are an efficient way to pick up information about the distribution of activity along the continuum. Together, different sensors are sensitive to information that is not available to individual sensors. In this way, the diverse response characteristics of the visual, vestibular, and somatosensory systems may be complementary with respect to complex patterns of self motion.

Efficient control of flight also requires that the pilot has some form of knowledge about the exchange of kinetic and potential energy (although this does not assume that the pilot has an "internal model" that is easily described by classical physics). An important basis for this knowledge is information about the ways in which changes in velocity are resisted in flight. Such information is contained in the relationship of control actions (e.g., stick, rudder, and throttle adjustments) to changes in aircraft states (e.g., velocity and trajectory). To the extent that one perceives the amplitude and frequency dependence of this relationship, the moment-to-moment dynamics of the aircraft are perceived. A more thorough understanding of the "nonstationary" dynamics of flight involves a sensitivity to the dependence of the dynamics on characteristics of the trajectory (e.g., G forces), the air mass (e.g., atmospheric pressure), and the aircraft (e.g., gross weight). This requires that the pilot frequently explore the relationship between control actions and aircraft states. Sensitivity to i.e., feedback about) control actions depends on characteristics of the controls (e.g., moveability of the stick). "Control loaders" are valuable in flight simulation because they allow the moveability of the

161

control stick to vary as a function of the simulated aerodynamic environment. however, the pick up of this information is dependent on the motion and force environment inside the cockpit (i.e., vibration and G magnitude to which the pilot is subjected). A challenge for developments in flight simulation is to design systems that provide information about the motion and force environment inside the cockpit.

1.3.3 Coordination of postural control and aircraft control. The focus here is on the fact that the pilot's body is not a single rigid structure attached rigidly to the aircraft. This has important consequences for perception and control whenever the velocity vector or attitude of the aircraft changes. Consider the effect on the pilot's body when the aircraft undergoes a linear acceleration or a change in attitude. Torques are produced in different ways in different parts of the body. These torques give rise to uncontrolled body movements unless they are resisted by muscular action (and, to some extent, by restraint devices in the cockpit). When the head moves relative to the cockpit, visual stimulation will not be specific to motion of the aircraft through the environment, and vestibular stimulation will not be specific to motion of the aircraft relative to an inertial reference frame. Stimulation of the somatosensory system (and to some extent, the visual system) will be specific to motion of the body relative to the cockpit. Note that multimodal stimulation is not redundant, it is complementary (cf., Riccio & Stoffregen, 1988, 1989; Stoffregen & Riccio, 1988, 1989a). The overall pattern of stimulation is specific to the acceleration event, and event in which motion of the aircraft and motion of the body cannot be considered independently. The event must be considered in its entirety because of the consequences for perception and control: imposed motion of the head can frustrate the pick up of optical information; imposed motion of the torso or arms can frustrate manipulation of the control stick. A challenge for developments in flight simulation is to design systems for which the nonrigidity of the pilot has consequences for perception and action.

Consider also the effects on the pilot's body when the aircraft moves along a curved trajectory. It is often desirable for the z-axis of the aircraft to be parallel to the G vector. When they are not parallel, the various segments of the pilot's body must be "tilted" with respect to the cockpit in order to maintain a state of balance. The direction of postural balance in the cockpit provides information about the attitude of the aircraft relative to the G vector. Vestibular and somatosensory systems are sensitive to this information (cf., Riccio, Martin, & Stoffregen, 1988; Riccio, 1989). Sensitivity to this information could help the pilot fine tune the maneuver (e.g., coordinating attitude and airspeed). Attention to the direction of balance is also important for postural control in the aircraft seat. The pilot must detect imbalance in various body parts and detect the relative orientation of the support surfaces used to maintain balance (cf., Stoffregen & Riccio, 1988). Postural control stabilizes the "platform" for the perception and action systems (Riccio & Stoffregen, 1988). Deficiencies in postural control could compromise perception and control of the aircraft maneuver.

Focused attention on the orientation of the body and the aircraft relative to the G vector could cause the pilot to loose orientation with respect to the terrain. The terrain generally will not be perpendicular to the G vector or the aircraft z-axis. Managing the orientation of the aircraft relative to the G vector and the terrain, and the orientation of the body relative to the G vector and the aircraft, would seem to be an important, albeit complex, component of skilled flight control. This skill cannot be acquired in a simulator that does not allow the relative orientations of aircraft, G vector, and terrain to be manipulated independently. "Motion platform" simulators allow these orientations to be manipulated independently. However, they do not allow rotation to be manipulated independently of

162

tilt with respect to the G vector. This is required for accurate simulation of curved trajectories. For example, the perception of rotation without a change in tilt is veridical during a coordinated turn. A challenge for developments in flight simulation is to design systems that allow the independent manipulation of rotation and the relative orientations of aircraft, G vector, and terrain.

Another important aspect of curved trajectories is variation in the magnitude of the G vector. Variation in G magnitude can be large enough to have significant physiological and biomechanical consequences (see Kron, et al., 1980). Many of these effects impose "hard" limits on perception and action. For example, "gray out" precludes peripheral vision; increases in the weight of the limbs may render movement impossible. The aircraft control problems that arise because of hard limits can be viewed as errors of omission; required control actions are precluded. However, even small variations in G magnitude change the environmental constraints on perception and action. Such constraints are "soft" in the sense that they do not necessarily preclude perception and action. They change the dynamics of body movement; that is, they change the muscular actions required to achieve a particular interaction with the environment. This can lead to control problems if the pilot does not have motor skills that are appropriate for the new dynamics. The aircraft control problems that arise because of soft constraints can be viewed as errors of commission; inappropriate control actions are induced. It is important to emphasize that learning to control an aircraft also involves learning to control the interaction of the body and the aircraft. The latter is probably a nontrivial component of piloting skills in many flight scenarios. Inappropriate skills may be acquired in a simulator that does not include the soft biomechanical constraints encountered in variable G maneuvers.

The inter-dependencies between postural and aircraft dynamics also influence the response to transients. for example, there are several ways in which the pilot can minimize the deleterious effects of changes in aircraft velocity or attitude. Muscular effort can be exerted in the direction opposite to the anticipated force due to aircraft motion. Alternatively, muscular co-contraction may stiffen the body sufficiently when forces cannot be anticipated. If neither of these strategies can be used, less massive parts of the body may be used to "take up slack" in the imposed motion. For example, eyes can move in such a way that fixation on a distant object can by maintained; the arms can move in such a way that the positions of the hands are maintained with respect to the controls. These skills of coordinated motion are important when the intent is to maintain posture (or fixation) and when the intent is to change posture (or fixation). For many flight scenarios, learning the inter-dependencies between postural and aircraft dynamics should be as important as learning the dynamics of the aircraft alone. Simulations may be seriously deficient if these inter-dependencies are not included. There is no reason to believe that fidelity of postural dynamics is any less important than fidelity of the "aero model" in flight simulation.

1.3.4. Multimodal perception and constraints on control. The issues that are most important in this ecological perspective on flight simulation have to do with the consequences of variation in the attitude and/or velocity vector of the aircraft. These consequences involve the forceful interaction of the aircraft with the pilot's body. For example, the forces imposed on the pilot's body stimulate multiple perceptual systems. It is a common assumption in many areas of research, including those concerned with flight simulation, that multimodal stimulation is either redundant or conflicting. However, this assumption is inappropriate given that nonredundancies are both common and informative for a nonrigid body (Riccio & Stoffregen, 1988, 1989; Stoffregen & Riccio, 1988, 1989a). Multimodal stimulation is more accurately described as complementary. The complementarity of

multimodal stimulation has nontrivial implications for simulator design. While redundant stimulation would be necessary if it provided information not available to individual perceptual systems.

The forces imposed on the pilot during flight not only change the stimulation of perceptual systems but also change the constraints on body posture and movement. Both imposed stimulation and biomechanical constraints provide information about the flight situation. The difference between these two sources of information is that sensitivity to the latter requires that the pilot is active in the cockpit. For example, head movements, arm movements, and balance reveal the dynamics of the environment in which they occur. The balance and movement of the head would seem to be particularly informative because of its multiplicity of motion sensors and because of its relative lack of support. It follows that control of the head should be an important consideration in flight simulation.

Stimulation in the aircraft and the simulator are different because the actual motion of the pilot and cockpit are different. A major design problem in flight simulation is that increasing the fidelity of some modes of stimulation often reduces the fidelity of other modes of stimulation. The designer must assess the relative importance of various modes of stimulation (e.g., particular devices and "drive algorithms") as sources of information (sometimes viewed as "cues"). Multimodal stimulation and constraints on control appear to complicate the process in the sense that more sources of information must be considered. However, they may actually simplify the process in that they provide additional criteria on which to assess the relative importance of various modes of stimulation. For example, a motion platform or a "helmet loader" (see Kron, et al., 1980) may increase fidelity of simulated acceleration with respect to the control of a nonrigid body (i.e., postural control), while a wide field-of-view visual display may reduce fidelity with respect to the same criteria.

Fidelity criteria that are based on postural control may require more justification than criteria that are based on aircraft control. This emphasizes the need for basic research on the issues mentioned above. However, there are other factors that may influence whether postural criteria will ultimately appear in flight simulation. For example, consider the problem of "simulator sickness." In spite of considerable interest in simulator sickness, there has been a notorious lack of progress in understanding this and other situations that induce "motion sickness" (Stoffregen & Riccio, 1989a). A recent theory of motion sickness argues that the malady is due to a prolonged interference with postural control (Riccio & Stoffregen, 1989). The theory accounts for a much greater range of nausogenic and non-nausogenic phenomena than do other theories. Stated simply for the case of simulator sickness: postural control will be disrupted in the simulator to the extent that it is based on simulated motion (e.g., optic flow) that is not related to the dynamics of balance in the simulator cockpit. It remains to be seen whether this theory will have any impact on the flight simulation community; however, there is increasing interest in postural control outside the simulator after adaptation to the simulator. any effect on postural control outside the simulator would have to explained in terms of the postural controls strategies acquired in the simulator. This would ultimately lead to an appreciation of the importance of postural control in the simulator.

1.4 Summary and Experimental Prologue

Flight simulation has been almost exclusively concerned with simulating the motions of the aircraft. Physically distinct subsystems are often combined to simulate the varieties of aircraft motion. "Visual display systems" simulate the motion of the aircraft relative to remote objects and surfaces

164

(e.g., other aircraft and the terrain). "Motion platform" simulators recreate aircraft motion relative to the gravitoinertial vector (i.e., correlated rotation and tilt as opposed to the "coordinated turn" in flight). "Control loaders" attempt to simulate the resistance of the aerodynamic medium to aircraft motion. However, there are few operational systems that attempt to simulate the motion of the pilot relative to the aircraft and the gravitoinertial vector. The design and use of all simulators is limited by poor understanding of postural control in the aircraft and its effect on the perception and control of flight. analysis of the perception and control of flight (real or simulated) must consider that (a) the pilot is not rigidly attached to the aircraft and (b) the pilot actively monitors and adjusts body orientation and configuration in the aircraft.

It was argued that this more complete approach to flight simulation requires that multimodal perception be considered as the rule rather than the exception. Moreover, the necessity of multimodal perception was revealed by emphasizing the complementarity rather than the redundancy among perceptual systems. The next sections outlines an experiment motivated by a workshop held recently at the NASA Ames Research Center (July, 1989). This experiment reflects some of the concerns mentioned above in that it considers possible consequences of coordination between postural and vehicular control.

2.0 PRELIMINARY EXPERIMENTAL DESIGN

2.1 Objective

In an exploratory experiment, we will evaluate predictions made by sensory-conflict and postural-instability theories of simulator sickness (cf. Riccio & Stoffregen, 1989; Stoffregen & Riccio, 1989). Experimental manipulations will be a comprise between operational relevance and theoretical relevance. Dependent variables will include "objective" measures of simulator sickness and its hypothetical correlates. In particular, we will evaluate the effects of our manipulations on several physiological measures of discomfort, several measures of postural control, and the experience of induced self motion (vection). The effects of the independent variables and the relationships among the dependent variables will be useful in the design and evaluation of flight simulators.

2.2 Apparatus

The experiment requires the use of a flight simulator in which discomfort and sickness are commonly reported. We plan to use the LHX helicopter simulator. This is a fixed-base simulator that has a wide (110 deg) field-of-view, high-resolution graphics, and a head-slaved helmet-mounted display. The display should contain objects on a textured terrain. In some conditions, the instrument panel inside the cockpit will be visible through a "window" in the outside-the-cockpit display. We will need to perturb the aircraft states with well-defined disturbances. The disturbances will be generated by a sum of three to seven harmonically unrelated sinusoids. The disturbance power will be concentrated in the frequency range between .01 and 1.0 Hz. A trial duration on the order of three to four minutes and a sampling rate of at least 60 Hz would be desirable. In some conditions, the pilot's head and torso will be restrained with an upper torso "seat belt" and shoulder harness. Demands on control of the head will be reduced with a cervical collar.

During a trial (not necessarily all trials), we will need to collect data on (a) the aircraft states that are relevant to the pilot's control task, (b) the pilot's flight-control actions, (c) the six degrees-of-freedom of head movement, and (d) physiological measures of discomfort (e.g., gastric motility and eye muscle activity).

We will also need to construct a zig-zag "balance beam" track to assess stability of gait outside the simulator.

2.3 Procedure

The simulated aircraft will move at a constant speed and altitude over a flat terrain. The aircraft will be subjected to a roll-axis disturbance. The first factor in the experimental design will be whether or not the pilot's head and torso are restrained. The second factor in the design will be task of the pilot. The task will be either (a) visually track an object that is not along the direction of motion (no control of the aircraft), (b) simply maintain the head and upper torso in an erect posture (no control of aircraft), or (c) disturbance regulation in which the pilot attempts to maintain a wings-level attitude.. The third factor in the experiment will be the presence or absence of an inside-the-cockpit scent. These factors will be manipulated in a fractional factorial design.

After each trial, pilots will rate the magnitude of vection that they experienced during the trial. A four-point rating scale will be used.

After a set of trials, the pilot will walk on a balance beam that curves alternately to the left and the right. The time to traverse the balance beam and the number of falls will be recorded.

We will also collect data on the pilot's subjective experience of discomfort. Pilots will be queried about symptoms ranging from eye strain and fatigue to nausea and dizziness.

2.4 Analyses

Physiological measures of discomfort will analyzed for each trial. The method of analysis vary from measure to measure. For example, the dominant frequency of gastric motility will be computed form the electrogastrogram (see Hettinger, et al., 1988). Subjective ratings of vection and discomfort will also be analyzed as in Hettinger, et al., 1988).

Manual control data will be analyzed for the disturbance regulation trials. We will compute the root-mean-square (RMS) roll-axis motion. We will compare the control-stick activity at the disturbance frequencies (correlated power spectrum) with the activity that is not at the disturbance frequencies (remnant power spectrum). We will compare the shapes of the correlated and remnant power spectra. We will compute the "open-loop" gain crossover frequency and phase margin. such analyses are generally informative in the disturbance regulation paradigm (Martin, McMillan, Warren, & Riccio, 1986; Riccio, Cress, & Johnson, 1987; cf. Zacharias, et al., 1986).

Head movement data will be analyzed on all trials. we will compute RMS activity for all degrees of freedom. We will compare the roll-axis head activity at the disturbance frequencies (correlated power spectrum) with the activity that is not at the disturbance frequencies (remnant power

spectrum). We will compare the shapes of the correlated and remnant power spectra for the roll axis. We also compute these frequency-domain statistics for any other axis for which there are differences in RMS head activity.

Sets of dependent variables will be analyzed by different investigators. There are five sets of dependent variables: (a) subjective measures of vection and discomfort, (b) physiological measures of discomfort, (c) manual control measures of disturbance regulation performance, (d) measures of postural stability in the simulator, and (e) measures of gait stability outside the simulator. The effects of the experimental manipulations on each set of dependent variables will be analyzed in separate analyses of variance. Individual analyses may be simplified by considering only subsets of the experimental manipulations. Collaboration among the investigators will facilitate analysis of the canonical correlations among the sets of dependent variables.

3.0 REFERENCES

Cardullo, F. M. & Sinacori, J. B. (1988). A broader view of in cockpit motion and force cuing. Paper presented at the 1988 Flight simulation Technologies Conference, Atlanta, GA

Flach, J., Riccio, G., McMillan, G., & Warren, R. (1986). Psychophysical methods for equating performance between alternative motion simulators. Ergonomics, 29, 1423-1438.

Hettinger, L. J., Kennedy, R. S., & Nolan, M. D. (1988, May). Vection and sickness in flight simulators. Paper presented at the Spring Meeting of the International Society for Ecological Psychology. Antioch College, Yellow Springs, OH.

Kron, G. J., Cardullo, F. M., & Young, L. R. (1980). Study and design of high G augmentation devices for flight simulators, (TR F-33615-77-C-0055). Binghamton, NY: Singer - Link Flight Simulation Division.

Lintern, G. (1987). Flight simulation motion systems revisited. Human Factors Society Bulletin, 30 (12).

Martin, E., McMillan, G., Warren, R., & Riccio, G. (1986). A program to investigate requirements for effective flight simulator displays. International Conference on Advances in Flight Simulation (23 pp.). London, England: Royal Aeronautical Society.

Riccio, G. E. (1989). Spatial orientation: How is orientation really perceived? (TR-8902). Champaign, IL: Gary E. Riccio Associates.

Riccio, G., & Cress, J. (1986). Frequency response of the visual system to simulated changes in altitude and its relationship to active control. Proceedings of the 22nd Annual Conference on Manual Control (17 pp.). Wright-Patterson AFB, OH: Aeronautical Systems Division.

Riccio, G., Cress, J., & Johnson, W. (1987). The effects of simulator delays on the acquisition of flight control skills: control of heading and altitude. Proceedings of the Human Factors Society - 31st Annual Meeting (pp. 1286-1290). New York, N.Y.: Human Factors Society.

Riccio, G. E., Martin, E. J., & Stoffregen, T. A. (1988). What's up? A new orientation. In F. E. McIntire (Ed.), Psychology in the Department of Defense, Proceedings of the Eleventh Symposium, (TR-88-1, pp. 326-330). Colorado Springs, CO: United States Air Force Academy.

Riccio, G. E., & Stoffregen, T. A. (1988). Affordances as constraints on the control of stance. Human Movement Science, 7, 265-300.

Riccio, G. E., & Stoffregen, T. A. (1989). An ecological theory of motion sickness and the affordances of instability. Part II of submitted manuscript.

Stoffregen, T. A., & Riccio, G. E. (1988). An ecological theory of orientation and the vestibular system. Psychological Review, 95, 3-14.

Stoffregen, T. A., & Riccio, G. E. (1989a). A critique of the sensory conflict theory of motion sickness. Part I of submitted manuscript.

Stoffregen, T. A., & Riccio, G. E. (1988b). Responses to optical looming in the retinal center and periphery. submitted.

Riccio, G. E., & Stoffregen, T. A. (1990). Gravitoinertial force versus the direction of balance in the perception and control of orientation. Psychological Review, in press.

Warren, R., & Riccio, G. (1986). Visual cue dominance hierarchies: implications for simulator design. SAE Transactions (15 pp.). Warrendale, PA: Society for Automotive Engineering.

Zacharias, G., Warren, R., & Riccio, G. (1986). Modeling the pilot's use of flight simulator visual cues in a terrain following task. Proceedings of the 22nd Annual Conference on Manual Control. Wright-Patterson AFB, OH: Aeronautical Systems Division.

GLOSSARY

Aerodynamics. The relationship between aircraft motion and the combined effects of commanded motion and changes in the air mass. to simplify the block diagrams, the automatic flight-control system and classical aerodynamics due to movements of the control surfaces and those due to changes in the air mass have not been differentiated.

Aero Disturbance. Changes in the air mass relative to the aircraft.

Aircraft (also a/c). An object that is capable of movement above ground through buoyancy or aerodynamics.

A/C Controls. The parts of the cockpit that can be moved or modified by the pilot in order to change or maintain the states of the aircraft.

A/C Visuals. Optical information from inside the cockpit: including the layout of surfaces in the cockpit as well as instruments.

A/C: Medium. Resistance of the medium of support (total aerodynamic environment) to particular aircraft states.

A/C: Object. States of the aircraft relative to another object.

A/C: Terrain. States of the aircraft relative to the ground.

Balance. Maintaining the orientation (or attitude) of a controlled system with respect to the vector sum of forces imposed on that system.

Biomechanics. The relationship between the motion of, and the total force acting on, various parts of an organism.

Coordination. Control of a part of an organism and/or its environment that takes into account the constraints imposed by concurrent control of another part of the organism and/or its environment.

Cost Functional. The effect of organismic and environmental parameters on the efficiency of action in a controlled system.

Disturbance. Changes in the states of aircraft relative to the terrain, other aircraft, or the air mass (including wind gusts).

Distal Layout. The parts of the substantial environment with which an organism is not in contact.

Environment. Surfaces of support (e.g., the terrain or the ground), media of support (e.g., an air mass or a non-contact force), detached objects (e.g., aircraft or projectiles), attached objects (e.g., trees or buildings).

Flight Simulator. A controlled system that recreates the motions and forces to which a pilot is subjected in an aircraft.

Flight control. A system that moves, or resists the movement of, the aircraft on the basis of information about the aircraft states (this is always the human in our block diagrams).

Gravitoinertial. The vector combination of gravity and acceleration, which can be conceptualized as an unitary force or as a potential for acceleration.

Imposed Forces. Vector combination of all forces acting on a particular part of an organism, excluding forces internal to the organism.

Manipulanda. The parts of the environment that can be moved or modified.

Medium. Parts of the environment that are nonsubstantial (i.e., afford passage through).

Object. Any substantial part of the environment that is distinct from the terrain or the ground (e.g., aircraft or projectiles).

Orientation of the Pilot. $\theta(t)$ and $\Phi(t)$.

Physiology. The systems internal to the organism that are effected by gravitoinertial magnitude.

Pilot: Balance. Orientation of various parts of the pilot's body (i.e., head, torso, arms, and legs) with respect to direction of balance.

Pilot: Controls. States of the pilot's manipulators (e.g., hands and feet) with respect to the a/c controls.

Pilot: Gravitoinertial Magnitude (also GI-mag). Physiological responses of the pilot to increases or decreases in the magnitude of the gravitoinertial vector.

Pilot: Seat. States of the pilot's body (i.e., torso and legs, including bottocks) with respect to a/c seat.

Pilot: Visuals. States of the pilot's eyes with respect to a/c visuals.

Postural Control. A system that, on the basis of information about body states, moves or resists the movement of the various parts of an organism that subserve balance.

Seat. Surface that can completely support the weight of the body through contact resistance at the buttocks, and that may resist the motion of the body through contact resistance at various parts of the torso and extremities (e.g., in an a/c seat).

Sensory Systems (also Perceptual System). Systems that can acquire information about states of an organism and its environment.

Self-Generated Forces. Forces internal to the organism that are responsible for moving, or resiting the movement of, parts of its body.

States of the Pilot/Aircraft. θ(t), Φ(t), ψ(t), x(t), y(t), z(t).

Terrain (also Ground). Surfaces that can completely support the weight f, and are large in scale relative to the action capabilities of, an object.

Vehicle. A controlled system that can transport an object form one place to another.

θ(t). Time history with respect to roll axis.

Φ(t). Time history with respect to pitch axis.

ψ(t). Time history with respect to yaw axis.

x(t). Time history with respect to longitudinal axis.

y(t). Time history with respect to lateral axis.
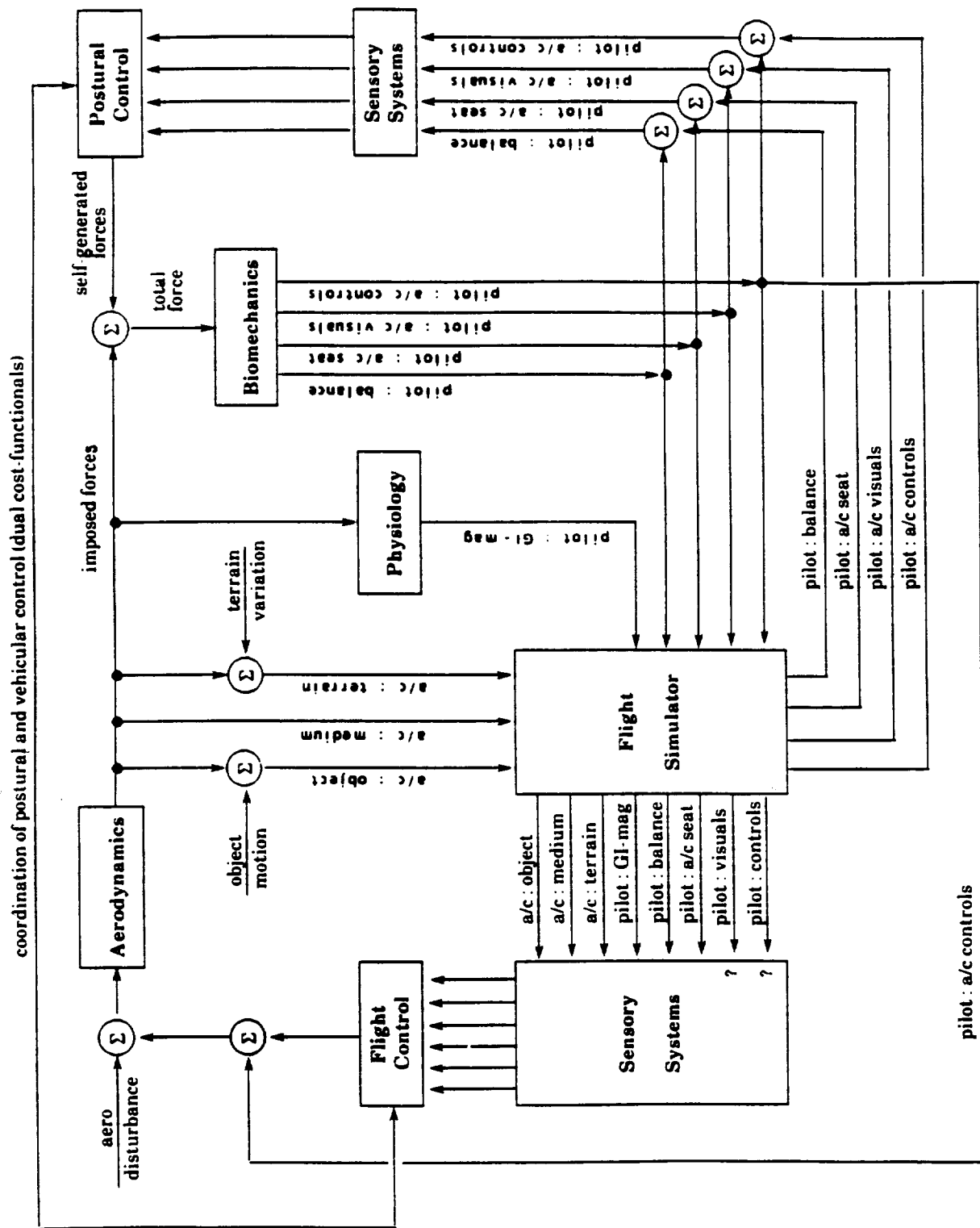
z(t). Time history with respect to gravity axis.
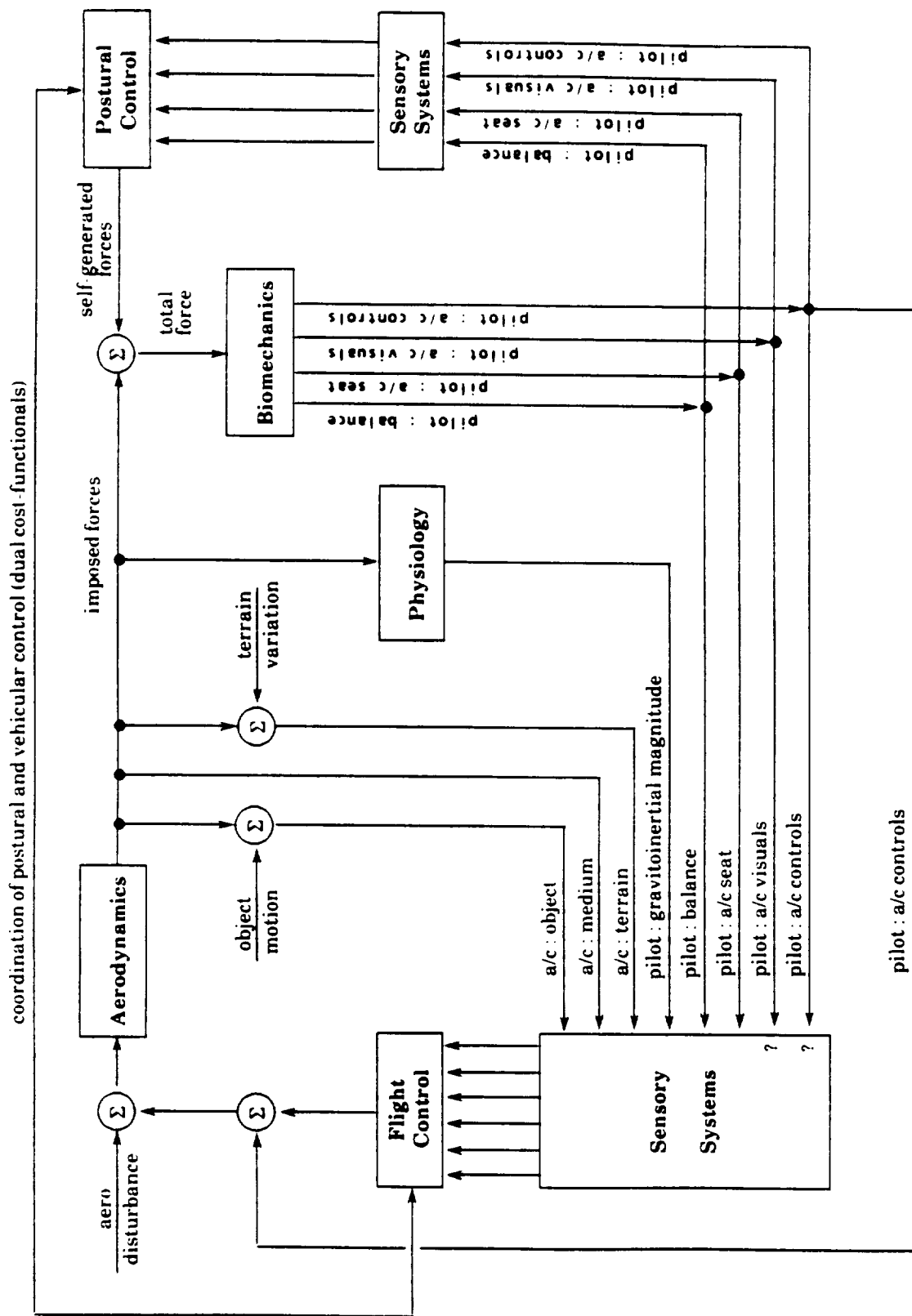
Figure 1. Coordination of postural control and vehicular control in an aircraft. The "closed loop" interactions between perception and control are indicated for the dual tasks of postural and aircraft control. Both control of posture and control of the aircraft require that the pilot monitor the "gravitoinertial" vector (i.e., the direction and magnitude of "G forces"), the parts of the environment that are in contact with the aircraft or the body, and some parts of the environment that are not in contact with the aircraft or the body. Note that control of the aircraft has consequences for control of posture and vice versa. Note also that the relationship between the aircraft and the pilot's body (e.g., relative position and orientation) provides information that is relevant to both postural control and aircraft control. See the glossary for a brief description of the labels used in the block diagram.

173

Figure 2. Coordination of postural control and vehicular control in a flight simulator. The "closed loop" interactions between perception and control are indicated for the dual tasks of postural and simulated aircraft control. Note the similarities and differences with Figure 1. The interactions between postural and aircraft control have important implications for simulator fidelity. The dynamics of postural control in the cockpit provide additional sources of information that could be exploited in flight simulation. See the glossary for a brief description of the labels used in the block diagram.

174

ILLUSORY SELF MOTION AND SIMULATOR SICKNESS

Lawrence J. Hettinger
Logicon Technical Services, Inc.
Dayton, Ohio

INTRODUCTION

According to the sensory conflict theory of motion sickness, spatially and/or temporally decorrelated perceptual information specifying one's dynamic orientation in space can lead to disorientation and sickness. The underlying conflict may either be intra- or intersensory in nature. Intrasensory conflict can arise, for instance, from decorrelated information within the vestibular system, such as that which accompanies Coriolis stimulation. Intersensory conflict can be caused by spatially and/or temporally decorrelated visual and vestibular information, such as that which occurs in flight simulators.

Simulator sickness is a form of motion sickness in which users of vehicular simulators exhibit signs and symptoms generally characteristic of motion sickness. In a fixed-base flight simulator, visual and vestibular sources of information specifying dynamic orientation are decorrelated to the extent that the optical flow pattern viewed by the "pilot" creates a compelling illusion of self motion which is not corroborated by vestibular information. Visually induced illusory self motion is known as "vection" (Tschermak, 1931) and a strict interpretation of sensory conflict theory makes vection in a fixed-base simulator a necessary precondition for simulator sickness.

This paper presents a discussion of simulator sickness (with applications to motion sickness and space sickness) based on the notion of the senses as perceptual systems (Gibson, 1966), and the sensory conflict theory (e.g., Reason & Brand, 1975). Most forms of the sensory conflict theory unnecessarily propose the existence of a "neural store." The neural store is thought to consist of a record of previous perceptual experiences against which currently experienced patterns of stimulation are compared. This paper seeks to establish that in its most parsimonious form the sensory conflict theory does not require a construct such as the neural store. In its simpler form, the sensory conflict theory complements and extends Gibson's view of the senses as perceptual systems.

I propose that motion and simulator sickness are produced by a breakdown (i.e., conflict) in the normal relationship between individual sub-systems of a functionally unitary perceptual system. The "orientation system," consisting primarily of the visual and vestibular sub-systems, is most directly implicated in the etiology of motion and simulator sickness. While in the case of simulator sickness illness may primarily be due to a breakdown on the stimulus side (i.e., decorrelated visual and vestibular information), in other cases disorientation and sickness can be produced by alterations in the normal activity of the physiological mechanisms that underlie the perception and maintenance of orientation (i.e., altered vestibulo-ocular reflex response in space sickness). Therefore a complete account of motion sickness, simulator sickness, and space sickness must address questions

concerning the "what" (i.e., the stimulus side) and the "how" (the neurophysiological side) of the phenomenon.

The sensory conflict theory also interacts well with most empirical and theoretical accounts of adaptation to perceptual distortion and perceptual learning. For instance, it is well known that, with time, humans and other animals adapt to the stimulus conditions that underlie motion sickness (Money, 1970), simulator sickness (Kennedy, Hettinger, & Lilienthal, 1990), and space sickness (Thornton, Moore, Pool, & Vanderpleg, 1987). Following adaptation to a nauseogenic force environment, readaptation to a previously benign force environment must occur and often results in a number of related perceptual-motor disturbances (e.g., land sickness, postural disequilibrium following simulator flights). Furthermore, the symptoms of disorientation, vertigo, mental confusion, and sickness that are characteristic of these maladies can be conceived as being due to a violation of normal multisensory relationships to which a lifetime of perceptual learning have made us uniquely sensitive.

The final section of this paper discusses a proposed experiment to be conducted on the U.S. Army's Crew Station Research and Design Facility at NASA Ames Research Center. The purpose of the experiment is to clarify the relationship between the experience of illusory self motion and the occurrence of simulator sickness, as well as to test the hypothesis that the onset of sickness in the simulator is preceded by a breakdown in the normal activity of postural control. This latter idea has been recently introduced by Stoffregen and Riccio (Personal Communication), and represents the first major new theoretical approach to motion sickness since the emergence of the sensory conflict theory.

SIMULATOR SICKNESS

This paper discusses the problem of simulator sickness, especially as it relates to the perception and control of self motion. A major purpose of the paper is to propose an experiment which could be conducted to clarify the relation between illusory self motion, or vection, postural instability, and simulator sickness.

Background

Motion sickness is a familiar, highly unpleasant condition which can occur when susceptible individuals are exposed to various provocative force environments, such as at sea, in space, in the air, and in vehicles on land. The capability to simulate aerial self motion has produced a new form of motion sickness referred to as "simulator sickness" (Kennedy, Hettinger, & Lilienthal, 1990; McCauley, 1984). Simulator sickness closely resembles "true" motion sickness (i.e., sea or air sickness), but is generally less severe and often involves visually-related disturbances (e.g., blurred vision, eyestrain) that are rarely observed in other forms of motion sickness (Ebenholtz, 1988).

Flight simulation has become an invaluable tool in the training and maintenance of aviator skills and in the research and development phases of aircraft design. This is due primarily to its inherent safety and cost effectiveness (Orlansky & String, 1977a; 1977b), as well as the wide range of

training and research scenarios that can be utilized. However, an apparent increase in the occurrence of simulator sickness threatens to diminish the utility of this technology for training and research and development.

Recent technical developments in flight simulation have stressed the use of large field-of-view visual displays of the out-of-the-cockpit scene using highly realistic imagery. The intent is to provide the user with a high degree of "felt presence" in the simulated environment. In parallel with, and possibly as a result of these technical developments, the reported incidence of discomfort, illness, and prolonged negative aftereffects among simulator users has steadily increased (Kennedy et al., 1990).

Simulator sickness may significantly limit the training and research capabilities of flight simulators. Illness is likely to have a negative effect on performance and learning, thereby contaminating research data and rendering training effectiveness questionable. When sickness is particularly frequent and severe, it may be necessary to restrict pilots' post-simulator flight activities, thereby diminishing their operational readiness. Pilot trainees may also adopt compensatory perceptual-motor strategies to avoid sickness in the simulator that will result in poor transfer of training to the aircraft. For example, pilots may restrict head movements in the simulator in order to minimize the occurrence of optokinetically-induced illness from pseudo-Coriolis effects (Dichgans & Brandt, 1973).

Symptoms of motion sickness are known to occur in the presence of visual stimulation alone with no concomitant physical movement (Dichgans & Brandt, 1978; Lestienne, Soechting, & Berthoz, 1977). Occurrences of illness while viewing Cinerama (Benfari, 1964) and other wide field-of-view motion displays (Parker, 1971) have been reported. For example, Lestienne, Soechting and Berthoz (1977) reported that subjects experienced intense, disturbing sensations of motion sickness induced by viewing large field-of-view, high velocity motion patterns. Three subjects out of thirty (10%) in their study became so disoriented while viewing these motion patterns that they fainted. The common element among these situations is the powerful, illusory sensation of self motion, referred to as "vection," experienced by the observers.

Vection, a term first used by Tschermak (1931), refers to the illusory sensation of self motion induced by viewing optical flow patterns that are specific to the form of self motion experienced. Vection can be induced in any of the body's linear or rotational axes (Dichgans & Brandt, 1978). Illusions of this sort are known to occur in non-laboratory conditions, such as the illusion of sudden forward motion induced by the perception of the backward motion of an adjacent automobile. Evoked responses in the vestibular nuclei of the rabbit (Dichgans & Brandt, 1972), cat (Daunton & Thomsen, 1976), and monkey (Henn, Young, & Finley, 1974) have been observed in response to vection-inducing stimuli, suggesting that such stimulation "recruits" activity in this area.

Until recently it was generally accepted that large field-of-view motion displays with substantial coverage of the peripheral retina were most effective in producing vection (Dichgans & Brandt, 1978). Andersen and Braunstein (1985), however, obtained reports of vection and motion sickness with centrally presented motion displays subtending visual angles as small as 7.5 deg. They asserted that an adequate representation of motion in depth may be as important as field-of-view size in

eliciting vection. Brandt, Wist and Dichgans (1975) obtained evidence indicating that the apparent motion of objects in depth is a powerful determiner of vection.

Reason and Brand (1975) hypothesized that in many cases conflicting inputs from visual and vestibular afferents are responsible for the occurrence of motion sickness. Intrasensory conflict (i.e., conflicting signals from the otoliths and semicircular canals) may also, in some cases, produce motion sickness. This "sensory conflict" theory would predict that visually induced apparent motion in the absence of corroborating vestibular motion information will produce motion sickness. To the extent that a visual stimulus depicts motion but does not also elicit vection, a conflict does not exist. Vection thus would appear to be a sine qua non for simulator sickness in fixed-base simulators according to this model. Individuals who report little or no illusory self motion in a fixed-base simulator should show little illness. The converse is not necessarily the case, because some individuals may be insensitive to such conflicts.

Previous experimentation on vection and sickness.

An experiment was recently conducted by Hettinger, Berbaum, Kennedy, & Nolan (in press) at the U.S. Navy's Visual Technology Research Simulator to investigate the relationship between vection and simulator sickness. Eighteen college student volunteers served as experimental observers. Each was asked to sit passively and observe three 15-minute computer generated representations of motion over a simulated 3-D terrain as presented on a large field of view (40 deg vertical, 80 deg horizontal) color visual display.

The motion trajectory presented to the observers was designed to be as nauseogenic as possible in order to assure that a sufficient number of observers experienced some symptoms of optokinetically-induced illness. It has been demonstrated that the most effective motion frequency for inducing sea sickness is slightly below 0.2 Hz (McCauley & Kennedy, 1976). Frequencies in this range were therefore selected for displacement in the vertical, longitudinal, and lateral axes, as well as for roll, pitch and yaw variations. All observers viewed the same motion patterns.

During the observation period, observers were asked to rate the degree of self motion they experienced on a scale of 0 - 3 (where 0 = "no feelings of self motion," 1 = "slight feelings of self motion," 2 = "moderate feelings of self motion," and 4 = "strong feelings of self motion"). Observers were also monitored for symptoms of simulator sickness using the Motion Sickness Questionnaire (Kennedy, McCauley, & Pepper, 1979) and also by means of an electrophysiological measure known as the electrogastrogram or EGG (Stern, etc.). The EGG measures the pacesetter potential of the stomach, which under normal circumstances is approximately 3 cycles/minute. When an individual becomes nauseated the EGG increases to a frequency of 11 - 15 cycles/minute,

The results indicated a clear and consistent relationship between the experience of illusory self motion and the occurrence of sickness. Those observers (approximately half) who reported no symptoms of sickness also reported little or no experience of illusory self motion. On the other hand, those observers who did experience sickness consistently reported moderate to strong sensations of self motion.

Visually-specified illusory self motion clearly represents a situation in which the normal activity of the orientation system is interrupted. Years of perceptual learning render most animals highly attuned to very specific temporal and spatial relationships between inputs from the visual and vestibular sub-systems. The coordinated, correlated activity of these sub-systems results in effective perception and maintenance of orientation and self motion.

Violation of these temporal and spatial constraints on the perception and maintenance of orientation appears to be the necessary prerequisite for disorientation and sickness. The evidence indicates that this is the case in flight simulators, in provocative terrestrial force environments, and in space sickness.

Discussion

Symptoms of motion sickness normally occur only in response to some form of physical displacement (e.g., motion at sea or in the air) with concomitant stimulation of the vestibular system. Therefore it may seem somewhat surprising to observe similar symptomatology in a fixed-base flight simulator in which no physical displacement occurs, but which may nonetheless provide compelling impressions of self motion.

The neural interrelationships between the visual and vestibular systems, primarily through the vestibular nuclei, have been the focus of a great deal of study in recent years (e.g., Precht, 1979). As I have argued above, it is generally useful to conceptualize visual and vestibular proprioception as manifestations of an integrated perceptual system (Gibson, 1966) designed to maintain orientation in space and control of self motion. Through a combination of heredity and a lifetime of perceptual learning, this system becomes attuned to spatial and temporal information which is highly correlated. The introduction of temporally asynchronous or distorted spatial information into the system appears to produce sensations of disorientation and illness in susceptible individuals

Simulations of in-flight visual motion patterns vary in the extent to which they elicit illusory sensations of self motion. Some provide veridical representations of optical flow patterns (Owen, 1982; Warren & Owen, 1982) characteristic of flight that do not lead to the illusion of self motion, while others appear to give rise to compelling experiences of vection. Researchers disagree on the requirements of visual displays for producing illusory self motion (e.g., Andersen & Braunstein, 1985). Nevertheless, the distinction between the perception of a depicted path (trajectory) and velocity of a point of view through a depicted space with no concomitant experience of illusory displacement, may be one of the keys to understanding the underlying causes of simulator sickness. Visually-specified, illusory self motion may entail a significant vestibular element while the perception of a display representing viewpoint motion without illusory displacement may not. A number of studies (e.g., Held, Dichgans, & Bauer, 1975; Mauritz, Dichgans, & Hufschmidt, 1977) have demonstrated large effects of rotating visual displays on postural sway. Observers in these studies continually readjusted their stance to compensate for visually-specified displacements of gravito-inertial upright. Lestienne et al. (1977) reported similar effects with patterns representing linear motion.

The relevance of these studies and the one reported here for the design of flight simulators lies in the demonstration that visual displays of motion patterns which produce vection produce more simulator sickness. In order to alleviate simulator sickness and related aftereffects it may be advantageous

C-3

to: (a) investigate the training utility of displays which do not produce illusory self motion, and/or (b) identify the underlying causes of sickness in displays that produce vection so they can be eliminated.

EXPERIMENTATION ON THE CREW STATION RESEARCH AND DESIGN FACILITY

Riccio and Stoffregen (1989) argue that simulator sickness, and other varieties of motion sickness, are due to prolonged interference with postural control. Their model states that: "Postural control will be disrupted in the simulator to the extent this it is based on simulated motion (e.g., optic flow) that is not related to the dynamics of balance in the simulator cockpit" (Riccio, 1989, p.12). The probability of sickness occurring in the simulator is therefore proportional to the amount of postural disruption.

Riccio and Stoffregen's model represents an opposing view to the sensory conflict theory. In particular, they object to the construct of the neural store which plays a central role in many versions of the conflict theory. Their model hypothesizes that a rather different form of conflict underlies the occurrence of disorientation and sickness. This conflict lies in the separate demands placed on strategies of postural control by the visual and somatosensory sub-systems of the orientation system.

By contrast, the version of the sensory conflict theory that I have argued for perceives the conflict to lie not at a motor control level, but at a somewhat more primitive sensory level. It is interesting to note that both models' predictions with regard to simulator sickness are best enhanced under a particular stimulus situation, i.e., with a highly effective (in terms of inducing sensations of self motion) visual depiction of self motion that has no corroborating somatosensory component.

The two models differ with regard to the predicted precursor signs of simulator sickness. The sensory conflict theory asserts that a powerful experience of the illusion of self motion is a necessary precondition for the occurrence of sickness in a fixed base simulator. The postural-instability model, on the other hand, would predict that sickness would be preceded by postural readjustments driven by the motion specified on the visual display. To the extent that postural readjustments are not observed (i.e., pilots' heads and torsos are restrained, or pilots simply do not respond to the visual display) sickness should not occur. Sensory conflict theory would predict that sickness would be largely independent of any postural control activity, although the experience of vection is often accompanied by postural control activity. We have endeavored to construct an experimental situation which would test the predictions of the two models.

Design

An exploratory experiment to evaluate these separate models of simulator sickness is proposed to be conducted on the U.S. Army's Crew Station Research and Design Facility (CSRDF) at NASA Ames Research Center. The CSRDF consists primarily of a fixed-base LHX helicopter simulator with a head-slaved helmet-mounted display that has a wide field-of-view (110 deg horizontal by 60 deg vertical).

During experimental trials the simulated aircraft will move at a constant speed and altitude over the simulated terrain. The terrain model in the CSRDF is produced using a General Electric Compuscene IV Computer Image Generation System, and provides a very realistic representation of highly textured terrain. During flight the aircraft will be subjected to roll-axis disturbance, generated by a sum of three to seven harmonically unrelated sinusoids. The disturbance power will be concentrated in the frequency range between .01 and 1.0 Hz.

In some conditions, the pilot's head and torso will be restrained to reduce demands on postural control. The torso will be restrained with an upper body harness, while the head will be restrained with the use of a cervical collar. Continuous ratings of the strength of illusory self motion will be obtained using either a verbal rating scale or a suitably rigged potentiometer.

Data will be collected during each trial on aircraft states, pilots' flight control actions, head movements, and physiological measures of discomfort. These latter measures include the electrogastrogram, electrocardiogram, blood volume pulse, respiration, skin temperature, skin conductance, and eye movement activity. Post-flight measures will include tests of postural equilibrium to assess ataxic effects of simulator exposure. Ataxia, or postural disequilibrium, is a common sign of simulator sickness (Kennedy et al., 1990).

The pilot's task will be to either visually track an object that is not along the direction of motion, maintain the head and upper torso in an erect posture, or maintain a straight and level attitude in the presence of the disturbance function. In the first two cases the pilot will have no control over the activity of the aircraft.

Five sets of dependent variables will be obtained: 1.) subjective measures of vection and discomfort, 2.) physiological measures of discomfort, 3.) manual control measures of disturbance regulation performance, 4.) measures of postural stability in the simulator, and 5.) measures of gait stability outside the simulator. Data analysis will concentrate on correlating magnitude estimates of vection and indices of postural control (i.e., head movement data) to our measures of sickness. Manual control data will also be analyzed for the disturbance regulation trials.

REFERENCES

Andersen, G.J. & Braunstein, M.L. (1985). Induced self-motion in central vision. Journal of Experimental Psychology: Human Perception and Performance, 11, 122-132.

Benfari, R.C. (1964). Perceptual vertigo: A dimensional study. Perceptual and Motor Skills, 18, 633-639.

Brandt, T., Wist, E.R., & Dichgans, J. (1975). Foreground and background in dynamic spatial orientation. Perception and Psychophysics, 17, 497-503.

Daunton, N.G. & Thomsen, D.D. (1976). Otolith-visual interactions in single units of cat vestibular nuclei. Neuroscience Abstracts II, 2, 1526.

Dichgans, J. & Brandt, T. (1972). Visual-vestibular interaction and motion perception. In J. Dichgans & E. Brizzi (Eds.), Cerebral control of eye movements and motion perception. New York: S. Karger.

Dichgans, J. & Brandt, T. (1973). Optokinetic motion sickness and pseudo-Coriolis effects induced by moving visual stimuli. Acta Otolaryngologica, 76, 339-348.

Dichgans, J. & Brandt, T. (1978). Visual-vestibular interaction: Effects on self-motion perception and postural control. In R. Held, H.W. Leibowitz, & H.L. Teuber (Eds.), Handbook of Sensory Psychology, Volume VIII: Perception. Berlin: Springer-Verlag.

Ebenholtz, S.M. (1988). Sources of asthenopia in Navy flight simulators (Final Rep. for Contract No. DAAL03-86-D-001). Orlando, FL: Naval Training Systems Center.

Gibson, J.J. (1966). The senses considered as perceptual systems. Boston: Houghton-Mifflin.

Held, R., Dichgans, J., & Bauer, J. (1975). Characteristics of moving visual scenes influencing spatial orientation. Vision Research, 15, 357-365.

Henn, V., Young, L.R., & Finley, C. (1974). Vestibular nucleus units in alert monkeys are also influenced by moving visual fields. Brain Research, 71, 144-149.

Kennedy, R.S., Hettinger, L.J., & Lilienthal, M.G. (1990). Simulator sickness. In G. Crampton (Ed.), Motion and space sickness. Boca Raton, FL: CRC Press, Inc.

Lestienne, F., Soechting, J., & Berthoz, A. (1977). Postural readjustments induced by linear motion of visual scenes. Experimental Brain Research, 28, 363-384.

Mauritz, K.H., Dichgans, J., & Hufschmidt, A. (1977). The angle of visual roll determines displacement of subjective visual vertical. Perception & Psychophysics, 22, 557-562.

McCauley, M.E. (Ed.) (1984). Simulator sickness: Proceedings of a workshop. Washington, DC: National Academy Press.

McCauley, M.E. & Kennedy, R.S. (1976). Recommended human exposure limits for very-low-frequency vibration (Rep. No. TR-76-36). Point Mugu, CA: Pacific Missile Test Center.

Orlansky, J. & String, J. (1977a). Cost-effectiveness of flight simulators for military training. Volume I: Use and effectiveness of flight simulators (IDA Paper P-1275). Alexandria, VA: Institute for Defense Analyses Science and Technology Division.

Orlansky, J. & String, J. (1977b). Cost-effectiveness of flight simulators for military training. Volume II: Estimating costs of training in simulators and aircraft (IDA Paper P-1275). Alexandria, VA: Institute for Defense Analyses Science and Technology Division.

Owen, D.H. (1982). Optical flow and texture variables useful in simulating self motion: II (Final Tech. Rep. for Grant No. AFOSR-81-0078). Columbus, OH: The Ohio State University, Department of Psychology, Aviation Psychology Laboratory.

Parker, D.M. (1971). A psychophysical test for motion sickness susceptibility. Journal of General Psychology, 85, 87-92.

Precht, W. (1979). Vestibular mechanisms. Annual review of neurosciences, 2, 265-289.

Reason, J.T., & Brand J.J. (1975). Motion sickness. London: Academic Press.

Riccio, G.E. (1989). Visually guided control of movement in the context of multimodal stimulation (Rep. No. TR-8904). Champaign, IL: Gary E. Riccio Associates.

Riccio, G.E. & Stoffregen, T.A. (1989). An ecological theory of motion sickness and the affordances of instability. Part II of submitted manuscript.

Thornton, W.E., Moore, T.P., Pool, S.L., and Vanderploeg, J.M. (1987). Clinical characterization and etiology of space motion sickness. Aviation, space and environmental medicine, 58(9, Suppl.), A1-8.

Tschermak, A. (1931). Optischer Raumsinn. In A. Bethe, G. Bergmann, G. Emden, A. Ellinger (Eds.), Handbuch der normalen und pathologischen physiologie. Berlin: Springer-Verlag.

Warren, R. & Owen, D.H. (1982). Functional optical invariants: A new methodology for aviation research. Aviation, Space, and Environmental Medicine, 53, 977-983.

Wiker, S.F., Kennedy, R.S., McCauley, M.E., & Pepper, R.L. (1979). Reliability, validity and application of an improved scale for assessment of motion sickness severity (USCG Tech. Rep. No. CG-D-29-79). Washington, DC: U.S. Coast Guard Office of Research and Development.

SPATIAL VISION WITHIN EGOCENTRIC AND EXOCENTRIC FRAMES OF REFERENCE

Ian P. Howard
York University
Toronto, Ontario

INTRODUCTION

It is remarkable that we are able to perceive a stable visual world and judge the directions, orientations and movements of visual objects given that images move on the retina, the eyes move in the head, the head moves on the body and the body moves in space. An understanding of the mechanisms underlying perceptual stability and spatial judgements requires precise definitions of relevant coordinate systems. An egocentric frame of reference is defined with respect to some part of the observer. There are four principal egocentric frames of reference, a station-point frame associated with the nodal point of the eye, an retinocentric frame associated with the retina, a headcentric frame associated with the head, and a bodycentric frame (torsocentric) associated with the torso. Additional egocentric frames can be defined with respect to any segment of the body. An egocentric task is one in which the position, orientation or motion of an object is judged with respect to an egocentric frame of reference. A proprioceptive task is a special kind of egocentric task in which the object being judged is also part of the body. An example of a proprioceptive task is that of directing the gaze toward the seen or unseen toe. An exocentric frame of reference is external to the observer. Geographical coordinates and the direction of gravity are examples of exocentric frames of reference. These various frames of reference are listed in Table 1, together with examples of judgements of each type.

The Station-Point Frame

We start with an illuminated three-dimensional scene of fixed objects, the visual world. A station point is defined with respect to some arbitrary coordinate system anchored in the world. Any optical system has two nodal points which have the geometrical property that all light rays passing through the first emerge from the second without having changed direction. The nodal points of the human eye are close together and can be regarded as one nodal point situated near the centre of the eye. The nodal point is a geometrical abstraction, light rays do not necessarily pass through it. The nodal point of the eye is the visual station point.

The visual surroundings or ambient array is the set of light sources and reflecting surfaces which surround the station point and from which light rays can reach the station point. The spherical array of light rays that reach the station point constitutes the station-point frame of reference. Within this frame of reference the distance of any point in the ambient array from the station point and the angle subtended at the nodal point by any pair of points in the ambient array can be specified. The station-point frame of reference itself contains no natural fiducial lines for specifying orientation or direction, since a point has no defined orientation. It is therefore meaningless to talk about the effects of

rotating the station point or rotating the ambient array round the station point. Every linear motion of the station point changes the distances of points and the angular subtense of pairs of points in the ambient array. The ambient array is sometimes thought of as the projection of the ambient array onto a fixed surface, usually a spherical surface centered on nodal point. This simply means that distances to points in the visual surroundings are not directly specified in this form of the ambient array, although the ambient array may contain enough information to allow distances to be recovered.

Visual attributes which may be defined in terms of the station-point frame of reference include (1) what is in view from a given place, (2) the relative directions (angular subtense) of two or more objects, (3) the distance of an object, (4) the relative angular velocities of moving objects, (5) velocity flow fields created by linear motion of the station point, expressed as a set of differential directed angular velocities and (6) the set of objects which define the locus of zero parallax (heading direction) in a three-dimensional array of objects as the station point is moved along a linear path. Judgements of these attributes are station- point judgements.

The Retinocentric Frame

We now add a pupil, lens, retina and associated structures of an eye. For a given position and orientation of the eye, the bundle of light rays which enter the pupil is the optic array and the portion of the ambient array from which these light rays originate is the distal visual stimulus, or field of view. For most purposes we can assume that the optic array projects onto a spherical retina centered on the nodal point. A visual line is any line which passes through the pupil and nodal point from a point in the distal stimulus to its image on the retina. The visual axis is the visual line through the fixation point and the centre of the fovea. A three-dimensional polar coordinate system centered on the nodal point can be used to specify the retinocentric distance, position, and direction of any object. An object's distance is its distance from the nodal point. Its eccentricity is the angle between its visual line and the visual axis. Its meridional direction is the angle between the plane containing the visual line of the object and the visual axis and the plane containing the visual axis and the retinal meridian which is vertical when the head is in a normal upright posture. These three-dimensional coordinates project onto the surface of the retina as two-dimensional polar coordinates, with the fovea as origin for eccentricity and the normally vertical meridian as the fiducial line for meridional direction. This is the retinocentric frame of reference. Note that the linear velocity of an image is proportional to the angular velocity of the object relative to the eye. This retinal coordinate system may be projected through the nodal point onto the concave surface of a perimeter or onto a tangent screen, which allows one to specify the oculocentric eccentricity and meridional angle of a stimulus on a chart. For certain purposes it may be more convenient to specify retinocentric positions in terms of elevation and azimuth or longitude and latitude. For instance, longitude and latitude are useful when describing the retinal flow field created by linear motion over a flat surface because the flow vectors conform to lines of longitude. The visual axis provides a natural reference which allows one to specify the direction of gaze with respect to selected landmarks in the ambient array.

Visual attributes which may be defined in terms of the retinocentric frame of reference include: (1) the eccentricity and meridional direction of an object (its visual direction relative to the fovea and the prime retinal meridian), (2) the orientation of a line, relative to prime retinal meridian. These are absolute retinocentric visual features. Relative retinocentric features involve only the specification of the relative positions, orientations or motions of images on the retina. Example are (1) the shape of a

186

retinal image (2) the retinal velocity of an image and hence the angular velocity of an object moving with respect to the eye. (3) the angular velocity of the eye with respect to a stationary object, and (4) retinal flow fields created by translations of the eye with respect to an ambient array. For a perfectly spherical retina relative retinocentric features are geometrically equivalent to those defined in terms of the ambient array projected onto a spherical surface.

All absolute retinocentric attributes change when the eye rotates with respect to a fixed nodal point and distal display. However, absolute retinocentric features are not necessarily affected by all types of eye rotation. For instance, the retinocentric direction of an object is invariant when the eye rotates about the visual line of the object. The eyes rotate as if about an axis at right angles to the meridian along which the gaze moves (Listing's law). An interesting consequence of this fact is that the retinocentric orientation a line is invariant when the gaze moves along the line (Howard, 1982, p. 185). For a spherical retina and distortion-free optical system, relative oculocentric attributes, such as the shape of the retinal image, are not affected by any rotations of the eye. If the retina were not spherical this would not be true and the task of shape perception would be more complex.

The station-point and retinocentric frames of reference are both oculocentric frames of reference.

The Headcentric Frame

We now add a head. The orientation of an eye in the head about each of three axes may be specified objectively in terms of either the Fick (latitude and longitude), the Helmholtz (elevation and azimuth) or the Listing (polar) coordinate system (see Howard, 1982 for details). The headcentric position of a visual object may be specified in terms of angles of elevation relative to a transverse plane through the eyes and angles of azimuth relative to the median plane of the head. The headcentric orientation of an object is usually specified with respect to the the normally vertical axis of the head. The head is defined as being vertical when the line from the ear hole to the angle of the eye socket and the line joining the two pupils are both horizontal. Particular headcentric spatial features of objects may be defined in terms of the types of head motion that leave them unchanged. If we assume that the centre of rotation of the eye is the same as the nodal point then the headcentric position of an object is the vector sum its retinocentric position and the position of the eye in the head. For instance, if an object is 10° to the left of the fixation point and the eye is elevated 10° then the headcentric position of the object is about 14.1° along the upper left diagonal with respect to the eye socket. Similar arguments apply to the headcentric orientation and motion of an object. Of course the coordinate systems used for specifying retinocentric position and eye position must correspond. Visual attributes that may be defined in headcentric terms include (1) the direction of an approaching object relative to the head, (2) the direction of gaze in the head, (3) an object's inclination to the mid-head axis and (4) a shape defined by the path an eye follows when pursuing a light spot.

The Bodycentric Frame

We now add a body. The bodycentric (torsocentric) position, orientation or movement of an object may be specified with reference to any of the three principal axes or planes planes of the body. The defining characteristic of bodycentric attributes is that they are affected by specific types of body motion.

If no part of the body is in view, bodycentric judgments require the observer to take account of oculocentric information, eye-in-head information and information from the neck joints and muscles regarding the position of the head on the body. Thus the oculocentric, headcentric and bodycentric reference systems form a hierarchical, or nested set, as indicated in the second column of Table 1. But this is not all. For certain types of bodycentric judgement the observer must appreciate the lengths of body parts, in addition to their angular positions. For instance, a person can place the finger tip of the hidden hand on a visual target only if the length of the arm is taken into account. Conscious knowledge is not involved, but rather the implicit knowledge of the body that is denoted by the term body schema. If the body as well as the object being judged is in view, bodycentric judgments are much simpler since they can be done on a purely visual basis without the need to know the positions of the eyes or head.

Examples of bodycentric attributes include 1) the direction of an object relative to a part of the body. This would need to be appreciated by a person who wished to direct the hidden hand towards an object, 2) motions of an object with respect to a part of the body and 3) the inclination of an object relative to the mid-body axis.

The Exocentric Frame

Finally, the exocentric position, orientation or movement of an object are specified with respect to coordinates external to the body. The defining characteristic of exocentric spatial attributes is that they are not affected by changes in the position or orientation of the observer or any part of the observer. Exocentric attributes may be absolute or relative. Absolute exocentric attributes are defined with respect to a coordinate system which is assumed to be fixed in inertial space. Examples of extrinsic coordinate systems are the one-dimensional gravitational coordinate, the two-dimensional geographical coordinates and a set of three-dimensional Cartesian coordinates. Absolute exocentric attributes include (1) the gravitational orientation of a line, (2) the compass direction of an arrow and (3) the movements of an object within a defined space.

Relative exocentric attributes are defined in terms of the position, orientation or motion of one object relative to another or of parts of an object relative to other parts. The reference frame is now intrinsic to the object or set of objects being judged. The distinction is analogous to that between extrinsic and intrinsic geometries. Relative exocentric attributes include (1) the shape of an object (the relative dispositions of parts), (2) rotation of an object relative to an intrinsic axis. For instance, the rotation of an aircraft about the yaw, roll or pitch axis and (3) the motion of one object relative to another.

Exocentric judgements about an isolated visual object can be with respect to a frame of reference provided by memory, as when we relate the position of a light to the remembered positions of the contents of a room. Otherwise, the exocentric position of an isolated visual object can be specified with respect to a frame of reference supplied by a second sense organ. Thus we can judge the position of a light in relation to a frame of reference provided by sounds or by things we touch or we can judge the orientation of a line in terms of stimulation registered by the vestibular organs. These are all intersensory tasks.

188

In theory, the variance of performance on an intersensory task should equal the sum of the variances of directional tasks that involve the separate component senses. A multisensory task is one in which the position, orientation or movement of an object is detected by more than one sense organ at the same time. For instance, we perform a multisensory task when we determine the headcentric direction of an object both by sight and by the sound that it makes. Given that the observer believes that the seen and heard object is one, the variance of performance on a multisensory task should, theoretically, be less the variance of performance on tasks using only one or other of the component senses (see Howard, 1982, Chapter, 11 for more details on the distinction between intersensory and multisensory tasks).

Finally there are cases where the frame of reference is external but the object is the self. I shall refer to them as semi-exocentric frames of reference. Examples of semi-exocentric attributes include (1) the position of an observer on a map (2) the compass direction of an observer with respect to an object (3) the position of an observer with respect to being under or over something. Note that, unlike purely exocentric attributes, semi-exocentric attributes vary with changes in the location of the observer.

In what follows I shall discuss the extent to which perceptual judgements within egocentric and exocentric frames of reference are subject to illusory disturbances and long-term modifications. I shall argue that well-known spatial illusions, such as the oculogyral illusion and induced visual motion have usually been discussed without proper attention being paid to the frame of reference within which they occur, and that this has lead to the construction of inadequate theories and inappropriate procedures for testing them.

PERCEPTUAL JUDGEMENTS WITHIN THE OCULOCENTRIC FRAME

The subjective registration of the station-point or retinocentric features of an object depend on the local sign mechanism of the visual system. This is the mechanism whereby, for a given position of the eye, each region of the visual field has a unique and stable mapping onto the retina and visual cortex.

Any misperception of the oculocentric position or movement of a visual object can arise only as a result of some disturbance of the retinal local sign-system or of the oculocentric motion-detecting system. In a geometrical illusion, lines are apparently distorted or displaced when seen in the context of a larger pattern. In a figural aftereffect a visual test object seen in the neighborhood of a previously seen inspection object appears displaced away from the position of the inspection object. Such effects operate only over distances of about one degree of visual angle and the apparent displacement rarely exceeds a visual angle of a few minutes of arc (Kohler and Wallach, 1944). We must conclude that the local-sign system is relatively immutable. This is not surprising, since the system depends basically on the anatomy of the visual pathways. Several claims have been made that oculocentric distortions of visual space can be induced by pointing with hidden hand to visual targets seen through displacing prisms (Cohen, 1966; Held and Rekosh, 1963). Others have claimed that these effects were artifactual and we are left with no convincing evidence that oculocentric shifts can be induced in this way (see Howard, 1982, page 501 for a more detailed discussion of this subject).

The movement aftereffect is a well known example of what is almost certainly an oculocentric disturbance of the perception of motion. I will not discuss this topic here.

PERCEPTUAL JUDGEMENTS WITHIN THE HEADCENTRIC FRAME

A person making headcentric visual judgements must take account of both oculocentric and eye-in-head information. The question of how and to what extent people make accurate use of eye-in-head information when making headcentric judgements is a complex one. One complication arises because the two eyes are in different positions. The visual system must construct a headcentric frame of reference that is common to both eyes. It can be shown that people judge the headcentric directions of an object as if the eyes were superimposed in the median plane of the head, somewhere between the actual positions of the two eyes. This is known as the cyclopean eye, or visual egocentre (See Howard, 1982 for a fuller discussion of all these issues).

A misjudgment of the headcentric direction or motion of a visual object can arise from a misregistration of the position or motion of either the retinal image or the eyes. In this section I shall consider only phenomena due to misregistration of the position or movement of the eyes.

Illusory Shifts of Headcentric Visual Direction

Deviations of the apparent straight ahead due to misregistered eye position are easy to demonstrate. If the eyes are held in an eccentric position a visual target must be displaced several degrees in the direction of the eccentric gaze to be perceived as straight ahead. When the observer attempts to look straight ahead after holding the eyes off to one side, the gaze is displaced several degrees in the direction of the previous eye deviation. Attempts to point to visual targets with unseen hand are displaced in the opposite direction. The magnitude of these deviations has been shown to depend on the duration of eye deviation and to be a linear function of the eccentricity of gaze (Hill, 1972; Morgan, 1978; Paap and Ebenholtz, 1976). Similar deviations of bodycentric visual direction occur during and after holding the head in an eccentric posture (Howard and Anstis, 1974). It has never been settled whether these effects are due to changes in afference or to changes in efference associated with holding the eyes in a given posture. Whatever the cause of these effects, it is evident that the headcentric system is more labile than the oculocentric system. This is what one would expect because headcentric tasks require the neural integration of information from more than one sense organ.

The Oculogyral Illusion

The oculogyral illusion may be defined as the apparent movement of a visual object induced by stimulation of the semicircular canals of the vestibular system (Graybiel and Hupp, 1946). The best visual object is a small point of light in dark surroundings and fixed with respect to the head. When the vestibular organs are stimulated, as for instance by accelerating the body about the mid- body axis, the point of light appears to race in the direction of body rotation. The oculogyral illusion also occurs when the body is stationary but the vestibular organs signal that it is turning. This happens, for instance, in the 20 or 30 seconds after the body has been brought to rest after being rotated. It is not surprising that a point of light attached to the body should appear to move in space when the

observer feels that the body is rotating. I shall refer to this perceived motion of the light with the body as the exocentric component of the oculogyral illusion. The exocentric component is not very interesting because it is difficult to see how a rotating person could do other than perceive a light which is attached to the body as moving in space. But even casual observation of the oculogyral illusion reveals that the light appears to move with respect to the 1 0 head in the direction of body acceleration. This is the headcentric component of the oculogyral illusion.

Whiteside et al. (1965) proposed that the headcentric component of the oculogyral illusion is due to the effects of unregistered efference associated with the vestibulo-ocular response (VOR). The idea is that when the subject fixates the point of light, VOR engendered by body acceleration is inhibited by voluntary innervation. The voluntary innervation is fully registered by the perceptual system but the VOR efference is not, and this asymmetry in registered efference causes the subject to perceive the eyes as moving in the direction of body rotation. This misperception of the movement of the eyes is interpreted by the subject as a headcentric movement of the fixated light. To support this theory we need evidence that the efference associated with VOR is not fully registered by the perceptual system responsible for making judgments about the headcentric movement of visual objects.

For frequencies of sinusoidal head rotation up to about 0.5 Hz, the vestibulo-ocular reflex (VOR) is almost totally inhibited if the attention is directed to a visual object fixed with respect to the head (Benson and Barnes, 1978). The most obvious theory is that VOR suppression by a stationary object is due to cancellation of the VOR by an equal and opposite smooth pursuit generated by the retinal slip signal arising from the stationary light. This cannot be the whole story because Barr et al. (1976) reported that the gain of VOR produced by sinusoidal body rotations decreased to about 0.4 when subjects imagined that they were looking at an object rotating with them. It looks as though VOR efference can be at least partially cancelled or switched off even without the aid of visual error signals (McKinley and Peterson, 1985; Melvill Jones et al. 1984). Tomlinson and Robinson (1981) were concerned to account for how an imaginary object can inhibit VOR but for our present purposes, the more important point is that VOR is not totally inhibited.

Perhaps an imagined object is not a satisfactory stimulus for revealing the extent of voluntary control over VOR. We wondered whether an afterimage might be a better stimulus because it relieves subjects of the task of imagining an object and requires them only to imagine that it is stationary with respect to the head. We had already found OKN to be totally inhibited by an afterimage even though it was not inhibited by an imaginary object. The results of all these experiments are reported by Howard et al. (1988).

Subjects in total darkness were subjected to a rotary acceleration of the whole body of $14°/s^2$ to a terminal velocity of $70°/s$, which was maintained for 60 s. In one condition subjects were asked to carry out mental arithmetic. In a second condition they were asked to imagine an object rotating with the body, and in a third condition an afterimage was impressed on both eyes just before the trial began and the subject was asked to imagine that it was moving with the body. The same set of conditions was repeated but with lights on so that the stationary OKN display filled the visual field. Under these conditions both VOR and OKN are evoked at the same time.

191

In all conditions the velocity of the slow phase of each nystagmic beat was plotted as a function of time from the instant that the body reached its steady-state velocity. For none of the subjects was VOR totally inhibited at any time during any of the trial periods. For the OKN plus VOR condition subjects could see a moving display, but they could totally inhibit the response only after about 30s, when the VOR signal had subsided.

We propose that VOR is not completely inhibited by an afterimage seen in the dark because the mechanism used to assess the headcentric motion of visual objects does not have full access to efference associated with VOR. Thus the system has no way of knowing when the eyes are stationary. The component of the VOR which cannot be inhibited by attending to an afterimage gives an estimate of the extent to which VOR efference is unregistered by the system responsible for generating voluntary eye movements and for giving rise to the headcentric component of the oculogyral illusion.

PERCEPTUAL JUDGEMENTS WITHIN THE EXOCENTRIC FRAME

Information about the position, orientation and movement of the body in inertial space is provided by the normally stationary visual surroundings, by proprioception and by the otolith organs and semicircular canals of the vestibular system. The otolith organs respond to the pitch and roll of the head with respect to gravity but provide no information about the rotation or position of the head around the vertical axis. The otolith organs also respond to linear acceleration of the body along each of three orthogonal axes but cannot distinguish between head tilt and linear acceleration. The semicircular canals provide information about body rotation in inertial space about each of three orthogonal axes. But if rotation is continued at a constant angular velocity the input from the canals soon ceases. The integral of the signal from the canals can provide information about the position of the body but only with respect to a remembered initial position.

Vection

Vection is an illusion of self motion induced by looking at a large moving display. For instance, illusory self rotation, or circularvection, is induced when an upright subject observes the inside of a large vertical cylinder rotating about the mid-body axis (yaw axis). For much of the time the cylinder seems to be stationary in exocentric space and the body feels as if it moving in a direction opposite to that of the visual display. Similar illusions of self motion may be induced by visual displays rotating about the visual axis (roll axis) or about an axis passing through the two ears (pitch axis). Judgements about the motion of the self with respect to an external frame of reference are semi-exocentric judgements since they involve an external frame and a reference to the self. Rotation of a natural scene with respect to the head is normally due to head rotation and the vestibular system is an unreliable indicator of self rotation except during and just after acceleration. Therefore it is not surprising that scene rotation is interpreted as self rotation, even when the body is not rotating. There is a conjunction of visual and vestibular inputs into the vestibular nuclei (Waespe and Henn, 1978) and the parietal cortex (Fredrickson and Schwarz, 1977) which probably explains why visual inputs can so closely mimic the effects of vestibular inputs.

Vection for different postures and axes of rotation

If the vection axis is vertical, the sensation of self rotation is continuous and usually at the full velocity of the stimulus motion. If the vection axis is horizontal, the illusory motion of the body is restrained by the absence of utricular inputs that would arise if the body were actually rotating. Under these circumstances a weakened but still continuous sensation of body rotation is accompanied by a paradoxical sensation that the body has tilted only through a certain angle (Held et al. 1975). There are three vection axes with respect to the body (yaw, roll and pitch) and in each case the vection axis can be either vertical or horizontal. Of these six stimulus conditions only three had been investigated. We decided to measure vection and illusory body tilt under all six conditions (Howard et al., 1987). The subject was suspended in various postures within a large sphere that could be rotated about a vertical or horizontal axis. The magnitude of vection and of illusory body tilt were measured for yaw, pitch and roll vection for both vertical and horizontal orientations of each axis (see Figure 1).

For body rotation about both vertical and horizontal axes, yaw vection was stronger than pitch vection which was stronger than roll vection. When the vection axis was vertical, sensations of body motion were continuous and usually at, or close to the full velocity of the rotating visual field. When the vection axis was horizontal, the sensations of body motion were still continuous but were reduced in magnitude. Also for vection about horizontal axes, sensations of continuous body motion were accompanied by sensations of illusory yaw, roll or pitch of the body away from the vertical posture. The mean illusory body tilt was about 20° but the body was often reported to have tilted by as much as 90°. Two subjects in a second experiment reported sensations of having rotated full circle. Held et al. reported a mean illusory body tilt of 14°. We obtained larger degrees of body tilt probably because our display filled the entire visual field and subjects were primed to expect that their bodies might really tilt. In most subjects, illusory backwards tilt accompanying by pitch vection was much stronger than illusory forward tilt . Only two of our 16 subjects showed the opposite asymmetry, that was also reported by Young et al. (1975).

Vection and the relative distances of competing displays

The more distant parts of a natural scene are less likely to rotate with a person than are nearer parts of a scene, so that the headcentric motion of more distant parts provides a more reliable indicator of self rotation than does motion of nearer objects. It follows that circularvection should be related to the motion of the more distant of two superimposed displays. In line with this expectation Brandt et al. (1975) found that vection was not affected by stationary objects in front of the moving display but was reduced when the objects were seen beyond the display. Depth was created by binocular disparity in this experiment and there is some doubt whether depth was the crucial factor as opposed to the perceived foreground-background relationships of the competing stimuli. Furthermore, the two elements of the display differed in size as well as distance.

Ohmi et al.(1987) conducted an experiment using a background cylindrical display of randomly placed dots which rotated around the subject, and a similar stationary display mounted on a transparent cylinder which could be set at various distances between the subject and the moving display. The absence of binocular cues to depth allowed the perceived depth order of the two displays to reverse spontaneously, even when they were .well separated in depth. Subjects were asked to focus

alternately on the near display and the far display while reporting the onset or offset of vection. They were also asked to report any apparent reversal of the depth order of the two displays, which was easy to notice because of a slight difference in their appearance.

In all cases vection was experienced whenever the display that was perceived as the more distant was moving and was never experienced whenever the display perceived as more distant was stationary. Thus circular vection is totally under the control of whichever of two similar displays is perceived as background. This dominance of the background display does not depend on depth cues, because circularvection is dominated by a display that appears more distant, even when it is nearer. We think that perceived distance is not the crucial property of that part of the scene interpreted as background. When subjects focused on the moving display, optokinetic pursuit movements of the eyes occurred, and when they focused on the stationary display, the eyes were stationary. But such a change in the plane of focus had no effect on whether or not vection was experienced, as long as the apparent depth order of the two displays did not change.

Thus sensations of self rotation are induced by those motion signals most reliably associated with actual body rotation, namely, signals arising from that part of the scene perceived as background. Vection sensations are not tied to depth cues, which makes sense because depth cues can be ambiguous. Furthermore, vection sensations are not tied to whether the eyes pursue one part of the scene or another, which also makes sense because it is headcentric visual motion that indicates self motion, and this is detected just as well by retinal image motion as by motion of the eyes.

Vection and the central-peripheral and near-far placement of stimuli

It has been reported that circularvection is much more effectively induced by a moving scene confined to the peripheral retina than by one confined to the central retina (Brandt et al. 1973). In these studies, the central retina was occluded by a dark disc which may have predisposed subjects to see the peripheral display as background and it may have been this rather than its peripheral position which caused it to induce strong vection. Similarly, when the stimulus was confined to the central retina subjects may have been predisposed to see it as a figure against a ground, which may have accounted for the weak vection evoked by it.

Howard and Heckmann (1989) conducted an experiment to test this idea. The apparatus is depicted in Figure 2. The subject sat at the center of a vertical cylinder covered with randomly arranged black opaque dots. A 54° by 44° square display of dots above the subject's head was reflected by a sheet of transparent plastic onto a matching black occluder in the center of the large display. The central display could be moved so that it appeared to be suspended 15cm in front of or 15cm beyond the peripheral display. In the latter position it appeared as if seen through a square hole. In some conditions, one of the displays moved from right to left or from left to right at 30°/s while the other was occluded. In other conditions both displays were visible but only one moved and in still other conditions, both displays moved? either in the same direction or in opposite directions. In each condition subjects looked at the center of the display and rated the direction and strength of circularvection.

The results are shown in Figure 3. They reveal that vection was driven better by the peripheral stimulus acting alone than by the central stimulus acting alone. Indeed it was driven just as well by a

moving peripheral display with the center black or visible and stationary as by a full-field display. However, vection was reduced when the central display moved in a direction opposite to that of the peripheral display. When the peripheral display was visible but stationary the direction of vection was determined by the central display but only when it was farther away than the surround. This result is understandable when we realize that this sort of stimulation is produced, for example, when an observer looks out of the window of a moving vehicle. The moving field seen through the window indicates that the viewer is carried along with the part of the scene surrounding the window on the inside. When the moving central display was nearer than the stationary surround, a small amount of vection was evident in the same direction as the motion of the central display. We believe that the motion of the center induced apparent motion in the stationary surround, which in turn caused vection. We call this 'induced-motion vection. These experiments are a confirmation and extension of experiments conducted by Howard et al. (1987).

Induced Visual Motion, an Oculocentric, Headcentric and Exocentric Phenomenon

Induced visual motion occurs when one observes a small stationary object against a larger moving background and was first described in detail by Duncker (1929). For instance, the moon appears to move when seen through moving clouds. In a commonly studied form of induced motion the stationary object is seen within a frame which moves from side to side. In this stimulus configuration the moving frame changes in eccentricity and this may be responsible for some of the illusory motion of the stationary object. In order to study the effects of relative motion alone it is best to present the stationary object on a large moving background that either fills the visual field or remains within the confines of a stationary boundary.

We have evidence that induced visual motion occurs within the oculocentric, the headcentric and the exocentric system and that the mechanisms in the three cases are very different. As an oculocentric effect, it could be due to contrast between oculocentric motion-detectors. As a headcentric effect, it could be due to misregistration of eye movements. This could occur in the following way. Optokinetic nystagmus (OKN) induced by the moving background is inhibited by voluntary fixation on the stationary object. If the efference associated with OKN were not available to the perceptual system, but the efference associated with voluntary fixation were available, this should create an illusion of movement in a direction opposite to that of the background motion. This explanation, which I proposed in Howard (1982, p. 303) is analogous to that proposed by Whiteside et al. to account for the oculogyral illusion. It has been championed more recently by Post, and Leibowitz (1985), Post (1986) and Post and Heckmann (1986).

Induced visual motion can also be an exocentric illusion. It has been explained above that inspection of a large moving background induces an illusion of self motion accompanied by an impression that the background is not moving. A small object fixed with respect to the observer should appear to move with the observer and therefore to move with respect to the exocentric frame provided by the perceptually stationary background. This possibility was mentioned by Duncker.

We have recently devised psychophysical tests which can be used to dissociate the oculocentric, headcentric and exocentric forms of induced visual motion. These tests will now be described.

195

The key to measuring the oculocentric component of induced visual motion is to have two inducing displays moving in opposite directions, with a stationary test object on or near each. Nakayama and Tyler (1978) reported that a pair of parallel lines pulsing in and out in opposite directions induced an apparent pulsation of a pair of stationary lines placed between them. The apparent velocity of this induced motion was only about 0. 1°/s. This display is not ideal for measuring oculocentric induced visual motion since the outward and inward motion of the two induction lines mimics visual looming produced by forward body motion. An outwardly expanding textured surface is known to induce forward linear vection (Andersen and Braunstein, 1985; Ohmi and Howard, 1988).

A better stimulus for measuring oculocentric induced visual motion is that shown in Figure 4a. The two inducing stimuli move in a shearing fashion which does not mimic visual looming. If the gaze is directed at the boundary between the two moving displays, neither optokinetic nystagmus nor vection should occur. Any perceived relative motion between the two test spots must reflect oculocentric induced motion since headcentric or exocentric induced motion would affect the two objects in the same way. The task of judging the relative velocity of the test spots is simplified by using a procedure described by Wallach et al. (1978). The two test spots were moved vertically at a velocity of 2°/s with periodic fast returns and subjects estimated the apparent inclination of the path motion of one spot relative to that of the other spot. The apparent direction of motion of each spot is the resultant of its actual vertical motion and its apparent horizontal motion. With this display we have found the velocity of oculocentric induced motion to be about the same as that reported by Nakayama and Tyler.

The next step is to isolate the headcentric component of induced visual motion. Since the oculocentric component is confined to the region of the inducing stimulus, placing the test dot on a black band, as shown in Figure 4b, ensures that this form of induced motion will not occur. Again subjects judged the apparent slant of the path of a vertically moving spot, but this time pursuing it with the eyes. In a series of experiments we have shown that the apparent slant of the track is determined by headcentric induced motion and is not influenced by exocentric induced motion. This is probably because the frame of reference for judging the vertical is carried with the illusory motion of the body. The magnitude of headcentric induced motion was found to be about 2°/s, which is considerably larger than oculocentric induced motion (Heckmann and Howard, 1989; Post and Heckmann, 1986).

Finally we measured exocentric induced visual motion by having subjects estimate the velocity of illusory self motion induced by the motion of a large moving display. By definition this is a measure of the exocentric induced visual motion. People readily experience 100% vection at stimulus velocities of up to 60°/s and stationary visual objects appear to move in space at the same velocity as the apparent movement of the body. Thus exocentric induced visual motion can be many times larger than headcentric induced motion which in turn is several times larger than oculocentric visual motion.

The task of distinguishing between oculocentric, headcentric and exocentric components of any perceptual phenomenon and the task of discovering which sensory or cognitive processes may be responsible for a given phenomenon, require tests and procedures specifically designed for each case.

196

REFERENCES

Andersen, G. J. and Braunstein, M. L. (1985) Induced self-motion in central vision. J. Exp. Psychol., Hum. Percept. Perform., 11, 122-132.

Barr, C. C., Schulthies, L. W. and Robinson, D.A. (1976) Voluntary, non- visual control of the human vestibulo-ocular reflex. Acta Otolaryngol., 81, 365-375.

Benson, A. J. and Barnes, G. R. (1978) Vision during angular oscillation: The dynamic interaction of visual and vestibular mechanism. Aviat. space environ. Med., 49, 340-345.

Brandt, T., Dichgans, T. and Koenig, E. (1973) Differential effects of central versus peripheral vision on egocentric and exocentric motion perception. Exp. Brain Res., 16, 476-491.

Brandt, T., Wist, E. R. and Dichgans, J. (1975) Foreground and background in dynamic spatial orientation. Percept. Psychophys., 17, 497-503.

Cohen, H. B. (1966) Some critical factors in prism adaptation. Amer. J. Psychol., 79, 285-290.

Duncker, K. (1929) Uber induzierte Bewegung. Psychol Forsch., 22, 180-259.

Fredrickson, J. M. and Schwan, D. W. F. (1977) Vestibulo-cortical projection. In R. F. Naunton (Ed.) The Vestibular System. New York: Academic Press, pp 203-210.

Graybiel, A. and Hupp D. I. (1946) The oculo-gyral illusion. J. Aviat. Med, 17, 3-27.

Heckmann, T. and Howard, I. P. (1989) Induced visual motion: dissociation of exocentric and head-centric components. Invest. Ophthal. Vis. Sci., 30 (ARVO Abs.), 323.

Held, R., Dichgans, J. and Bauer, J. (1975) Characteristics of moving visual areas influencing spatial orientation. Vision Res., 15, 357-365.

Held, R. and Rekosh, J. (1963) Motor-sensoly feedback and geometry of visual space. Science, 141, 722-723.

Hill, A. L. (1972) Direction constancy. Percept. Psychophys., 11, 175-178.

Howard, I. P. (1982) Human Visual Orientation. Chichester; Wiley. pp 388-397.

Howard, I. P. and Anstis, T. (1974) Muscular and joint-receptor components in postural persistence. J. exp. Psychol., 103, 167-170.

Howard, I. P., Cheung, B. and Landolt, J. (1977) Influence of vection axis and body posture on visually-induced self rotation and tilt. Proceedings of the AGARD Conference on Motion Cues in Flight Simulation and Simulator Sickness. Brussels.

Howard, I. P., Giaschi, D. and Murasugi, C. M. (1988) Inhibition of optokinetic and vestibular nystagmus by head-fixed stimuli. Exp. Brain Res., Submitted for publication.

Howard, I. P. and Heckmann, T. (1989) Circular vection as a function of the relative sizes, distances and positions of two competing visual displays. Perception, in press.

Howard, I. P., Ohmi, M., Simpson, W. and Landolt, J. (1987) Vection and t'ne spatial disposition of competing moving displays. Proceedings of the AGARD Conference on Motion Cues in Flight Simulation and Simulator Sickness. Brussels.

Kohler, W. and Wallach. H. (1944) Figural aftereffects: An investigation of visual processes. Proc. Amer. Philos. Soc., 88, 269-357.

McKinley, P.A. and Peterson, B. W. (1985) Voluntary modulation of t'ne vestibuloocular reflex in humans and its relation to smooth pursuit. Exp. Brain Res., 60, 454-464.

Melvill Jones, G., Berthoz, A. and Segal, B. (1984) Adaptive modification of the vestibulo-ocular reflex by mental effort in darkness. Brain Res., 56, 149-153.

Morgan, C. L. (1978) Constancy of egocentric visual direction. Percept Psychophys., 23, 61-68.

Nakayama, K. and Tyler, C.W. (1978) Relative motion induced between stationary lines. Vision Res., 18, 1663-1668.

Ohmi, M., Howard, I. P. and Landolt, J. P. (1987) Circular vection as a function of foreground-background relationships. Perception, 16, 17-22.

Ohmi, M. and Howard, I. P. (1988) Effect of stationary objects on illusory forward self motion induced by a looming display. Perception, 17, 5-12.

Paap, K. R. and Ebenholtz, S. M. (1976) Perceptual consequences of potentiation in the extraocular muscles: An alternative explanation for adaptation to wedge prisms. J. exp. Psychol., 2, 457-468.

Post, R. B. (1986) Induced motion considered as a visually induced oculogyral illusion. Perception, 15,131-138.

Post, R. B. and Heckmann, T. (1986) Induced motion and apparent straight ahead during prolonged stimulation. Percept. Psychophys., 40: 263-270.

Post, R B. and Leibowitz, H. W. (1985) A revised analysis of the role of efference in motion perception. Perception, 14, 631-643.

Tomlinson, R. D. and Robinson, D.A. (1981) Is the vestibulo-ocular reflex cancelled by smooth pursuit? In Fuchs, A. and Becker, J. (Eds.) Progress in Oculomotor Research. Elsevier: North Holland.

Waespe, W. and Henn, V. (1978) Conflicting visual-vestibular stimulation and vestibular nucleus activity in alert monkeys. Exp. Brain Res, 33, 203-211.

Wallach, H., Bacon, J. and Schulman, P., 1978, Adaptation in motion perception: alteration of induced motion. Percept. Psychophys., 24: SO9-514.

Whiteside, T. C. D., Graybiel, A. and Niven, J. I. (1965) Visual illusions of movement. Brain, 88, 193-210.

Young, L. R., Oman, C. M. and Dichgans. J. M. (1975) Influence of head orientation on visually induced pitch and roll sensations. Aviat. space and environ. Med., 46, 246-269.

Table 1. Frames of Reference for Visual Spatial Judgements.
O signifies the object, the position or orientation of which is being judged or set
RF signifies the reference frame with respect to which the object is being judged or set

Frame of reference	Sensory components	Examples of tasks
Proprioceptive		
O and RF internal		
Non-visual	Sense of position of body parts	Point to the unseen toe
Purely visual	Locations of images of body parts	Align two seen parts of the body
Intersensory	Location of image plus proprioception	Point unseen finger to seen toe
Egocentric		
O external, RF internal		
Station point	Abstract or inferred	Specify objects visible from a vantage point
Retinocentric	Retinal local sign plus retinal landmark	Fixate an object. Place line on retinal meridian
Headcentric	Eye position + retinal local sign	Place an object in the median plane of the head
Bodycentric		
Purely visual	Relative retinal location	Align a stick to the seen toe
Intersensory	neck + eye position + retinal local sign	Point stick to the unseen toe. Place an object to left of the body
Semi-exocentric		
O internal, RF external		
Purely visual	Relative retinal location	Align self with two objects
Intersensory	Seen part of body and gravity senses	Point upwards
Exocentric		
O and RF external		
Absolute	Vision with appropriate reference frame	Judge geographical directions
Relative	Relative retinal location with appropriate constancies	Align three object. Judge the shape of an object
Intersensory	Visual and non-visual stimuli compared	Set a line vertical. point line to unseen sound

(a) Vertical yaw (b) Horizontal yaw

(c) Vertical pitch (d) Horizontal pitch

(e) Vertical roll (f) Horizontal roll
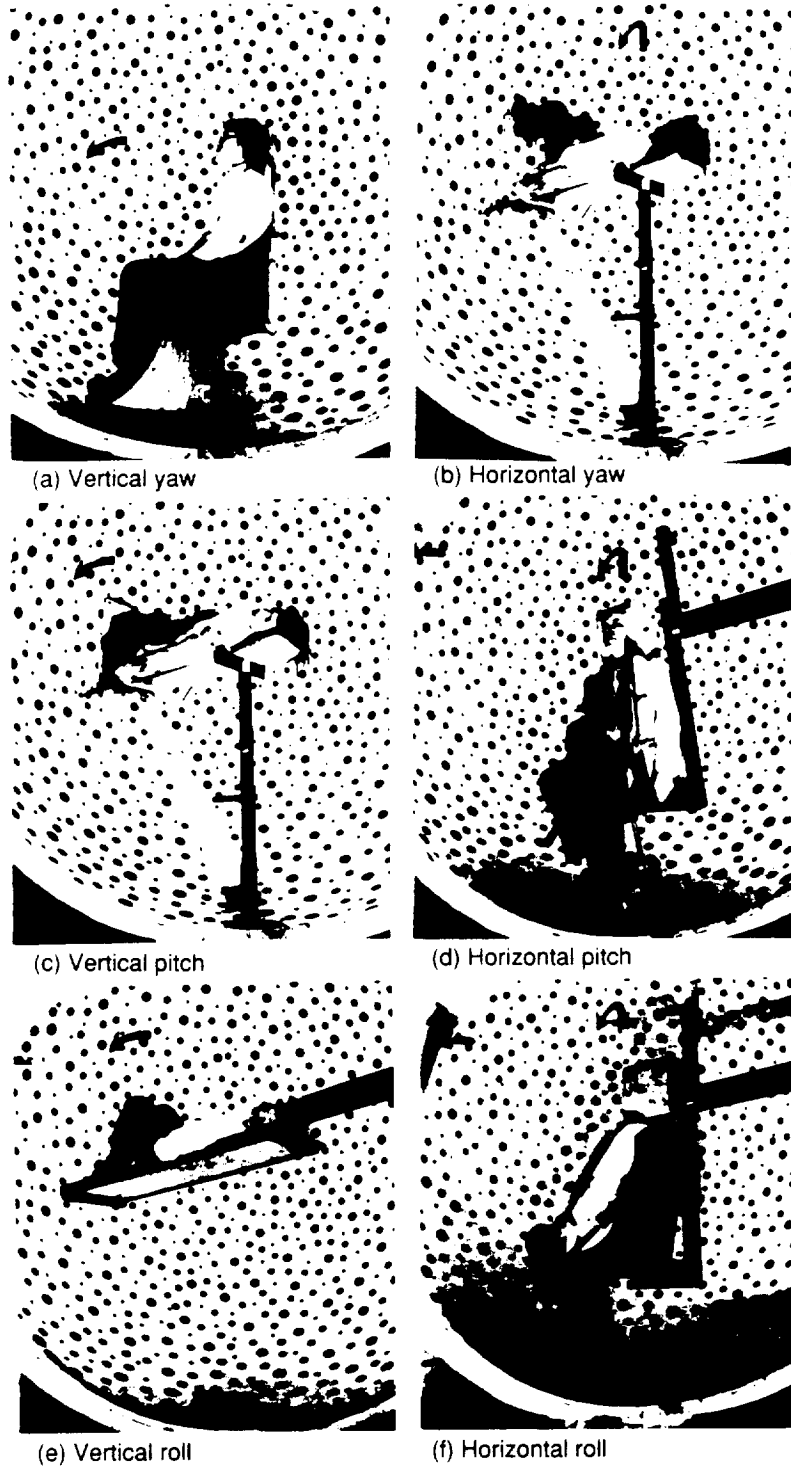
Figure 1. Stimulus conditions. Yaw denotes stimulus rotation about the mid-body axis, pitch about the y-body axis and roll about the visual axis. Vertical and horizontal refer to the orientation of the axis of scene rotation.
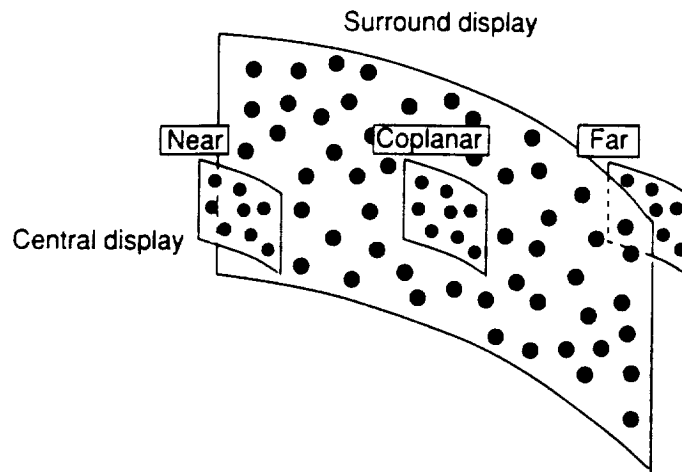
Figure 2. A diagrammatic representation of the displays used by Howard, Simpson and Landolt (1987) to study the interaction between central-peripheral and far-near placement of two displays in generating circularvection. The two displays could be moved in the same or in opposite directions, or one of them could be stationary or blacked out.
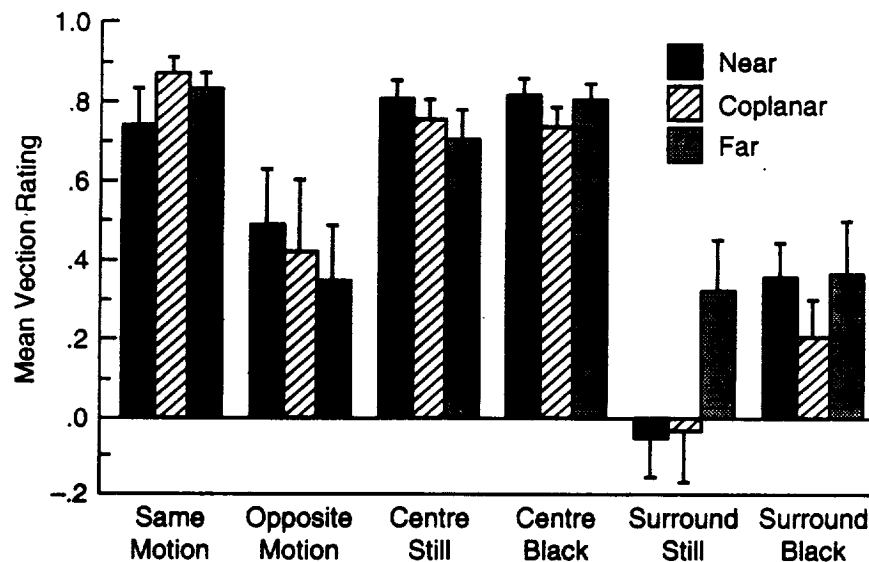
Figure 3. Mean vection ratings of nine subjects plotted as a function of the relative depth between the central and peripheral parts of the display and the type of display. A vection rating of 1.0 signifies full vection in a direction opposite to the motion of the display. When the two parts of the display moved in opposite directions, the motion of the peripheral part was taken a reference. The error bars are standard errors of the mean.
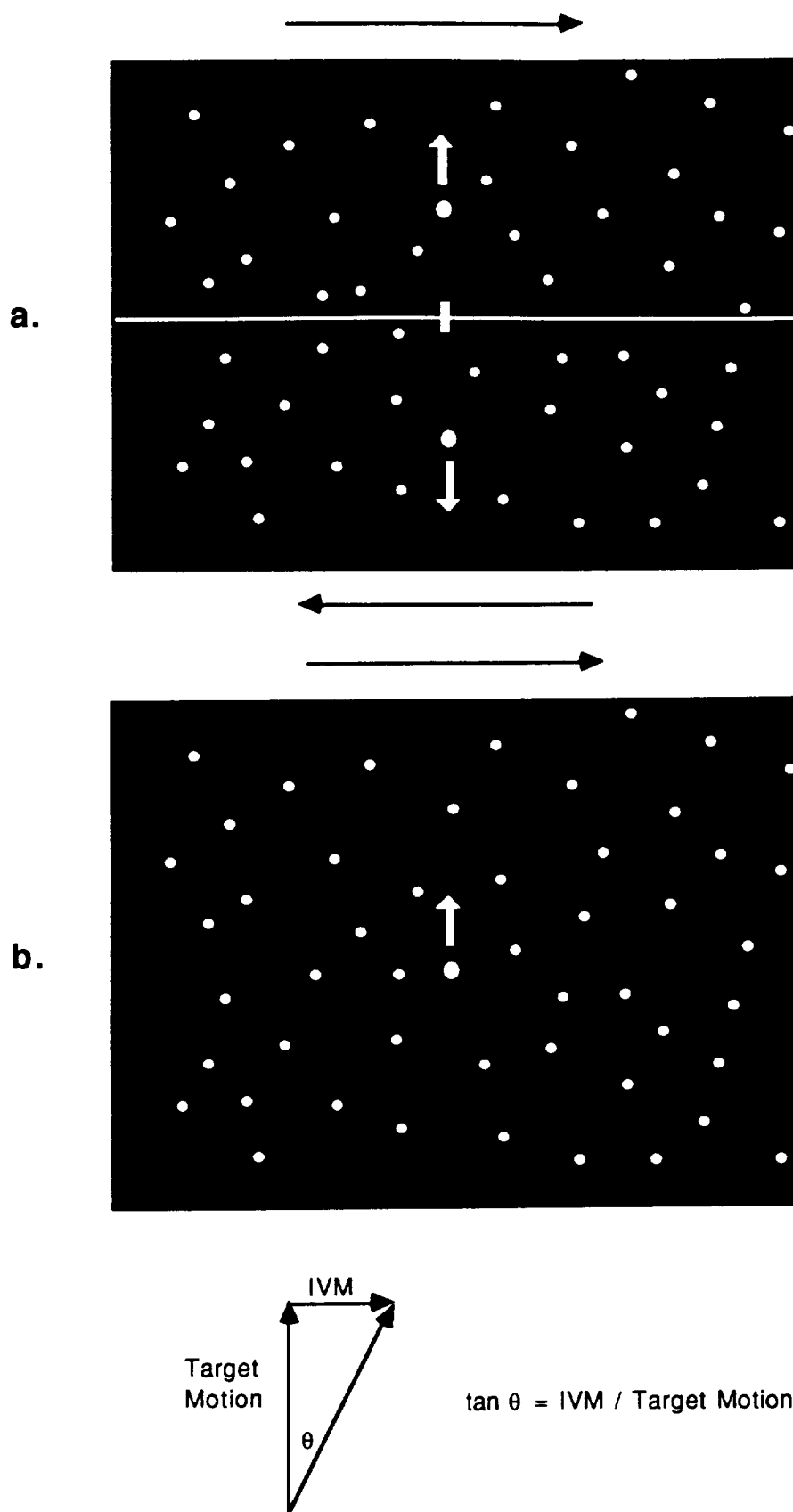
Figure 4. Stimuli for measuring components of induced visual motion: (a) Oculocentric component; (b) headcentric component.

VISUAL DIRECTION AS A METRIC OF VIRTUAL SPACE

Stephen R. Ellis,[1] Stephen Smith,[2] Selim Hacisalihzade[3]
NASA Ames Research Center
Moffett Field, California

ABSTRACT

Two experiments examine the abilities of 10 subjects to visualize directions shown on a perspective display. Subjects indicated their perceived directions by adjusting a head-mounted cursor to correspond to the direction depicted on the display. This task is required of telerobotic operators who use map-like pictures of their workspace to determine the direction of objects seen by direct view. Results show significant open-loop, judgement biases that may be composed of errors arising from misinterpretation of the map geometry and overestimation of gaze direction.

INTRODUCTION

A number of investigations and reviews of the characteristics of the virtual space perceived in pictures have been conducted recently (Rosinski et al.,1980; Sedgwick,1986; McGreevy and Ellis,1986; Grunwald and Ellis, 1986; Ellis, Smith and McGreevy,1987; Barfield, Sandford, and Foley, 1989). Despite the fact that the pictures considered were not stereoscopic, viewers typically were reported to develop a clear sense that the pictured objects were laid out in a virtual space. Quantitative characterization of the metrics of the viewer's perceived space will advance our understanding of picture perception and assist the design of displays for aircraft and spacecraft. The objective of the following research is to characterize patterns of errors observers make when referring a judged exocentric direction to a target presented on a perspective display to their own egocentric sense of visual direction. This type of spatial task is commonly faced by operators of telerobotic systems when using a map-like display of their workspace to determine the visual location and orientation of objects seen by direct view. It is also essentially the same task as faced by an aircraft pilot using a cockpit perspective traffic display of his surrounding airspace to locate traffic out his windows.

Previous studies of the error pattern in direction judgements have focused on exocentric judgements for which the subjects indicated their estimates of the target position by adjusting dials to show a target's azimuth and elevation with respect to a reference direction vector (See fig. 1).

---

[1] NASA Ames Research Center and U.C. Berkeley School of Optometry
[2] Sterling Software, Palo Alto, CA
[3] NRC Research Associate, NASA Ames Research Center

This response may be described as exocentric since the dial's frame of reference is external to the observer and contrasts with egocentric judgements in which target position is indicated with respect to a body-referenced coordinate system Accordingly, in order to test the generality of reported biases in estimating azimuth and elevation with exocentric judgements, it is useful to examine the same exocentric task but request the subjects to make egocentric judgements.

For this new response the observer adjusts the visual direction of head-mounted light cursor to indicate his sense of the target's depicted azimuth and elevation with respect to a reference position and reference direction. This response will explicitly test the generality of previously reported bias in which exocentric directions are judged to be away from a reference straight ahead. This bias may be attributed to errors in the subjects ability to determine the view direction used to generate the display (McGreevy and Ellis, 1986; Grunwald and Ellis, 1986; Ellis, Smith, Grunwald, and McGreevy, 1989). Furthermore, use of an egocentric response such as visual direction provides a more natural response than a dial adjustment. In a sense we ask the subjects to imagine themselves oriented in the virtual space along a particular direction vector and then to imagine where they would have to look to see the target.

METHODS

Two groups of 5 subjects participated as independent groups in two experiments. The subjects were male laboratory personnel ranging in age from 20 to 43 who were unfamiliar with the purpose of the experiment.

The experiments were conducted inside a 1.5 m planetarium dome that served as a projection surface for a head mounted, light pointer which projected a red filament image shaped as a 1.5° chevron onto the dome's surface (light from a 3v flashlight bulb through a Wratten #25 filter). The subject's head position was sensed by a Polhemus electromagnetic head tracker attached to an non-metallic modified welder's helmet approximately 11 cm above the head. The head tracker was independently calibrated against 28 theodolite-positioned, reference markers which were visible during calibration but not during testing.

The subjects were presented with an exocentric judgement task generated by a PDP 11/40 - Evans & Sutherland PS I graphics system. The images used were similar to earlier experiments (McGreevy and Ellis, 1986; Grunwald and Ellis, 1986; Ellis, Smith, Grunwald, and McGreevy, 1989). The major change was the greater yaw of the view direction used to create the images. It was set to a counterclockwise yaw of –35°. Pitch remained –22°. The subjects were seated at the center of the projection in front of the computer calligraphic monitor about 80 cm from the display surface and looked downward into it with a –22 deg. pitch angle matching that of the view vector. The viewport was 17 cm square.

Subjects were first positioned in an adjustable chair so that their head-mounted light cursor pointed to a subjective straight-ahead, eye level that corresponded to the calibration point at 0° pitch and 0° yaw. (See fig. 2) While in this position, a reference reading was taken from the head sensor for all future measurements. The subjects then were instructed to examine a series of automatically,

randomly presented displays and to estimate the azimuth and elevation direction of the target with respect to a reference position and direction. Then they were to transfer this judgement to their egocentric frame of reference. They made the judgement by adjusting the pitch and yaw of their head-mounted, light cursor to a position where they would expect to see the target if their head was at the reference position, and initially aligned with the reference direction in the displayed virtual space. For most of the judged directions the subjects could not simultaneously see the display and the cursor position, but had to gaze back and forth between them to accomplish the task, generally using head movements for excentricites greater than 15°. After adjusting the cursor, they held their position and moved a toggle switch that signaled the computer to take the data. The data for a 1 sec. period prior to the switch signal were averaged to give a single measurement. Three replications of each position were taken from each subject in a randomized sequence of 64 measurements that took about 2 hours to complete.

The interpretation of the head-direction data is complicated by the different centers of rotation associated with pitch, yaw and roll of the head. Pure yaws did not displace the center of rotation very much and the measured head yaw to the calibrated positions were within 2° of the calibrated angles within ±60° of the straight ahead, the greatest deviations being at the most extreme angles. The reason for the residual error was the difficulty of exactly positioning the subject to the calibration reference point. Pitch in contrast tends to be around a moving center of rotation somewhat behind the neck and consequently tends to translate the head upwards and backwards from the initial reference point which was used to provide a straight ahead, level reference for all subsequent measures. Consequently, when the subjects pointed their head-mounted cursors to the extreme pitches, the sensor reading undershot the calibrated value by from 5 to 8 degrees! We have calculated geometrical corrections for the effects of this displacement from the reference point since we could measure it, but generally found that they were small (2–4°) and for reasons discussed below may not in principle be proper to use.

After calibration of the head tracker in the light, the two experiments were conducted in the dark with the CRT display turned down so that only the frame of the monitor was faintly visible to provide an egocentric direction reference. *In one experiment the head cursor was kept on. In the other the cursor was turned off and the subjects had to rely principally on vestibular and proprioceptive cues to "look" to the direction they would expect to see the target.*

RESULTS AND DISCUSSION

The results from both experiments were similar and are analyzed together in this summary. Multivariate analysis of variance conducted with BMDP 4V on the elevation and azimuth errors showed that for both judgements the target elevation, target azimuth had statistically reliable effects on both the pitch and the yaw of the errors in the head pointing error. Pitch direction errors: Target Elevation: $F=16.14$ $df=4,5$ $p <.009$; Target Azimuth: $F=7.08$, $df=4,5$ $p<.027$; Yaw direction errors: Target Elevation: ns; Target Azimuth: $F=29.5$ $df=4,32$ $p<.001$. Standard errors for the mean error ranged between 1 and 10 degrees. The main effect of the presence of the light cursor was not significant and did not interact with other independent variables (See figs. 3 and 4)

207

Since we did not anticipate errors as large as actually measured, we did not use spherical statistics to correct the problems of mapping spherical data into a linear analysis. Because the analysis was conducted on the error data *corrected for wrap-around* of the scale and most of the errors were less than 15 degrees, use of spherical statistics is not likely to substantially change the major results.

The proper method to use to correct for movement of the subject from his calibrated reference position while he positions the cursor depends upon his interpretation of the of the meaning of the cursor position. If he considers its image on the inner surface of the sphere to represent the location of a target cube at that distance, about 1.5m, he would have to introduce parallax corrections to his body-referenced, head direction as he translated with respect to the original reference point so as to keep the cursor on the same place on the sphere as he moved. Alternatively, if he considered, as he in fact was instructed, the cursor image to represent a body-referenced direction toward the target, head displacement in itself would not require adjustment of head direction to keep to cursor properly pointed. This condition is particularly true since he was instructed that the target was at a relatively great distance from the reference cube. For the layouts used, the distance between target and reference was 6m and the viewing distance was modeled at 28m. At this distance the parallax correction for a 5 cm lateral movement would have to be only about 0.5°, comparable to the biological noise associated with head direction. Thus, since the head-angle was measured with respect to a body-referenced straight ahead, correction for head displacement need not be made.

The observed mean body-referenced errors for both experiments are plotted in figures 3 and 4 as error arcs on a rectangular projection of the response sphere. The pattern shows a tendency to err towards the subject's egocentric straight ahead, but with a significant asymmetry. The results may be interpreted as a composition of errors: 1) the asymmetrical pattern previously reported for exocentric dial responses which is generally away from the straight ahead and 2) a larger but symmetric tendency to overestimate the extent of the gaze direction indicated by the head mounted cursor. Over estimates like this have been reported by Biguer et al. (1984) for hand pointing to visual target and for head pointing to brief auditory targets (Perrott, Ambarsoom, and Tucker, 1987). In the case of hand pointing without visual feedback of pointing error such overestimates result in overshoot errors. In the case of head pointing without pointing error feedback, the overestimates result in undershoot errors similar to those observed.

The observation that the errors were not effected by turning off the light cursor supports the idea that one source of error arises from the proprioceptive and vestibular estimate of head rotation. But whether the phenomena is truly one of gaze remains to be determined by future experiments examining gaze angles produced by different combinations of eye and head angles. The results of the current study clearly show however, the visual direction is a significantly biased metric of virtual space presented by flat panel perspective displays. Modeling and explanation of the causes of the observed biases will allow design of compensated perspective displays.

REFERENCES

Barfield, W.; Sandford, J.; and Foley, J.: The mental rotation and perceived realism of computer generated three dimensional images. International Journal of Man-machine studies, 1989, (in press).

Biguer, B.; Prablanc, C.; Jeannerod, M.: The contribution of eye and head movements in hand pointing accuracy. Experimental Brain Research, vol. 55, 1984, pp. 462-469.

Grunwald, Arthur; and Ellis, Stephen R.: Spatial orientation by familiarity cues. Proceedings of the 6th European Annual Conference on Manual Control, June 1986, University of Keele, Great Britain.

Ellis, Stephen R.; Smith, Stephen; McGreevy, Michael W.: Distortions of perceived visual directions out of pictures. Perception and Psychophysics, vol. 42, 1987, pp. 535-544.

Mardia, K. V.: The statistics of directional data Academic Press, New York, 1972.

McGreevy, Michael W.; and Ellis, Stephen R.: The effect of perspective geometry on judged direction. Human Factors, vol. 28, 1986, pp. 439-456.

Perrott, David: Changes in head position as a measure of auditory localization performance: auditory psychomotor coordination under monaural and binaural listening conditions. Journal of the Acoustical Society of America, vol. 82, 1987, pp. 1637-1645.

Rosinski, R. R.; Mulholland, T.; Degelman, D.; and Faber, J.: Picture perception: an analysis of compensation. Perception and Psychophysics, vol. 28, 1980, pp. 512-525.

Sedgwick, H. A.: Space perception in Handbook of Perception and Human Performance vol. 1., K. R. Boff, L. Kaufman, J. P. Thomas eds, New York: Wiley, vol. 21, no. 19-21, 1986.
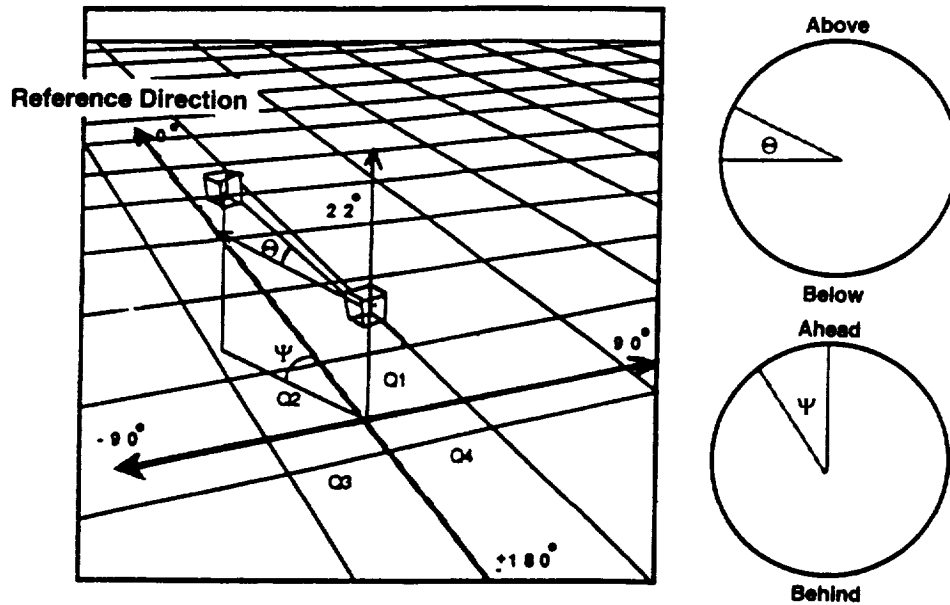
Figure 1. A schematic illustration of the direction judgement task. The subject adjusted the angles Ψ and Θ shown on the dials at the right until they appeared equal to the azimuth angle Ψ and the elevation angle Θ of the target cube relative to reference at the center. Dotted lines, labels and arrows did not appear on the map display.
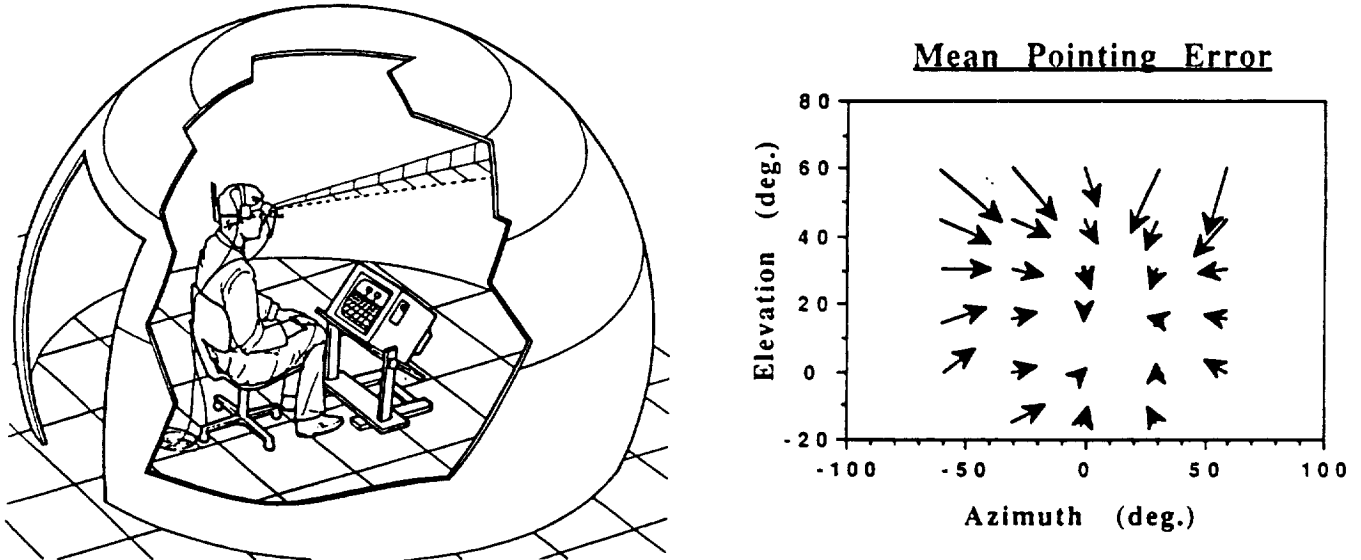
Figure 2. A schematic illustration of the experimental arrangement by which the subject indicated the visual direction at which he would expect to see the target presented on the CRT perspective display if he were positioned at the reference point and aligned with the reference direction. The data in the right portion of the figure represent the average error arcs in a rectangular projection of the forward sphere when both experimental conditions are combined. Each arrow represented the average pitch and yaw error in visual direction to a point at the tail of the arrow.

Mean Egocentric Direction Error
Curson On
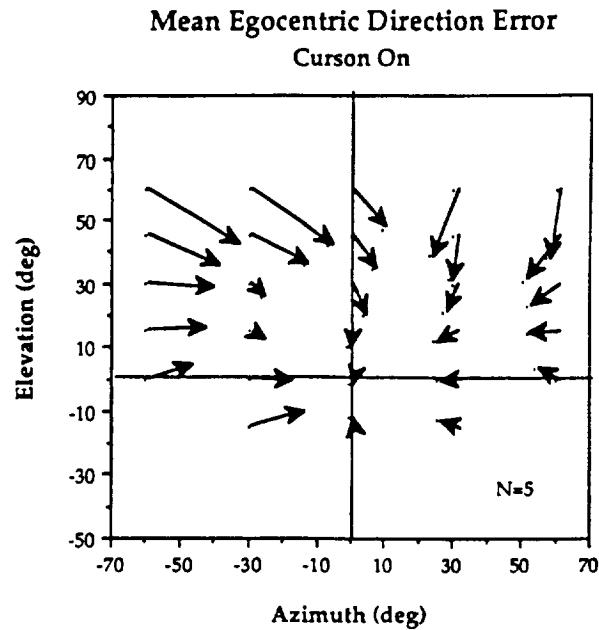
Figure 3. The data in the figure represent the average error arcs in a rectangular projection of the forward sphere for the condition in which the head driven cursor was turned on. Each arrow represents the average pitch and yaw error in visual direction to a point at the tail of the arrow.

Mean Egocentric Direction Error
Cursor Off
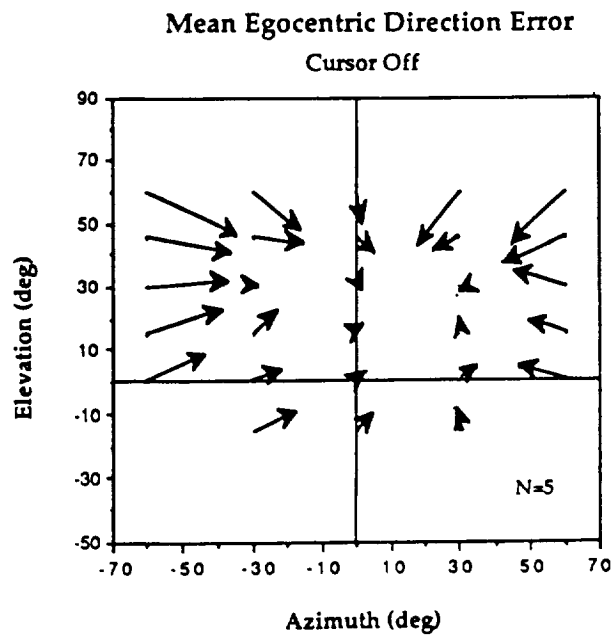
Figure 4. The data in the figure represent the average error arcs in a rectangular projection of the forward sphere for the condition in which the head driven cursor was turned off. Each arrow represents the average pitch and yaw error in visual direction to a point at the tail of the arrow.

211

Reference Direction

Eye
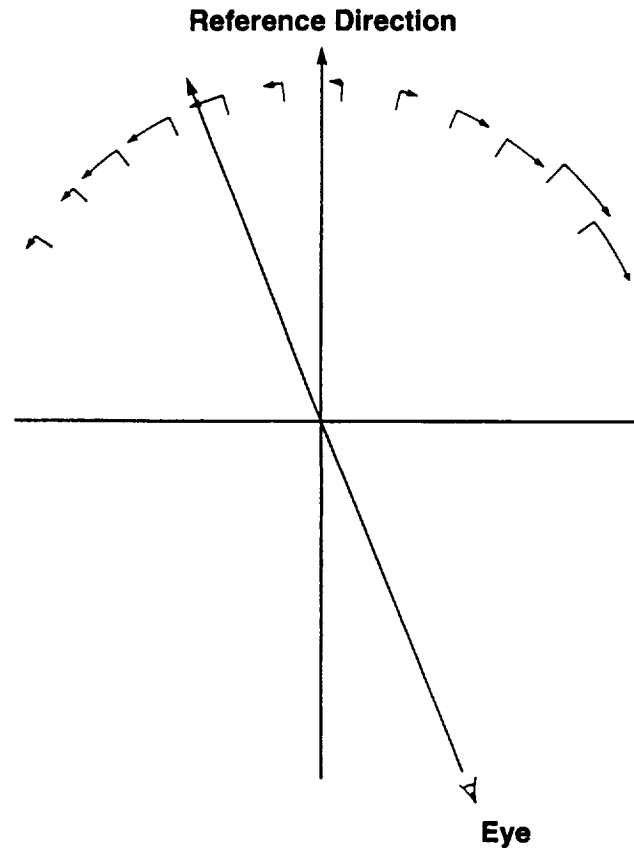
Figure 5. Circular plots for perspective displays in which subjects indicated target azimuth for targets at 0 degrees elevation by adjusting an angle on a dial. The errors are plotted as directed arcs with the tail of each arrow at the correct position of the target. The length of each arrow represents the average error from 8 subjects. Though the viewing azimuth was $-22°$ compared to the $-35°$ used in the current experiments, the conditions are otherwise comparable. The error arcs clearly show a bias away from the straight ahead rather than towards it and also show an asymmetry with greater errors in the right quadrant than in the left. Thus, if this bias were to cancel a larger one, perhaps due to overestimation of gaze direction, that was toward the straight ahead, the resulting bias would be smaller in the right quadrant than in the left. This expected pattern in found the the data for zero degree target orientation in figures 3 and 4.

APPENDIX A

PILOT/VEHICLE MODEL ANALYSIS OF VISUALLY GUIDED FLIGHT

Greg L. Zacharias
Charles River Analytics Inc.
Cambridge, Massachusetts

PILOT/VEHICLE MODEL ANALYSIS OF VISUALLY-GUIDED FLIGHT

o PILOT/VEHICLE MODEL DESCRIPTION

o CONTROL OF ALTITUDE WITH SIMPLE TERRAIN CUES

o SIMULATED FLIGHT WITH VISUAL SCENE DELAYS

o MODEL-BASED IN-COCKPIT DISPLAY DESIGN

o RANDOM THOUGHTS

OPTIMAL CONTROL MODEL OF PILOT/VEHICLE SYSTEM

VISUAL SUBMODELS

INSMOD
INSTRUMENT CUES

LINMOD
LINEAR PERSPECTIVE
CUES

TEXMOD
TEXTURE & FLOW-FIELD
CUES

VISUAL
CUEING
MODEL

216
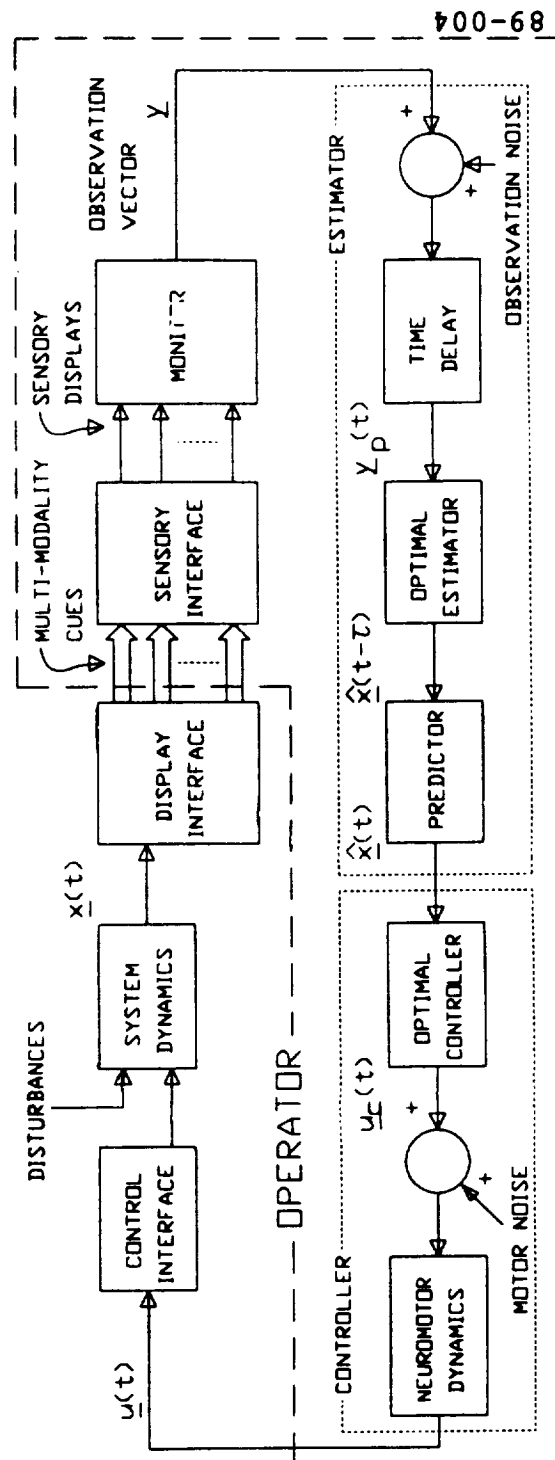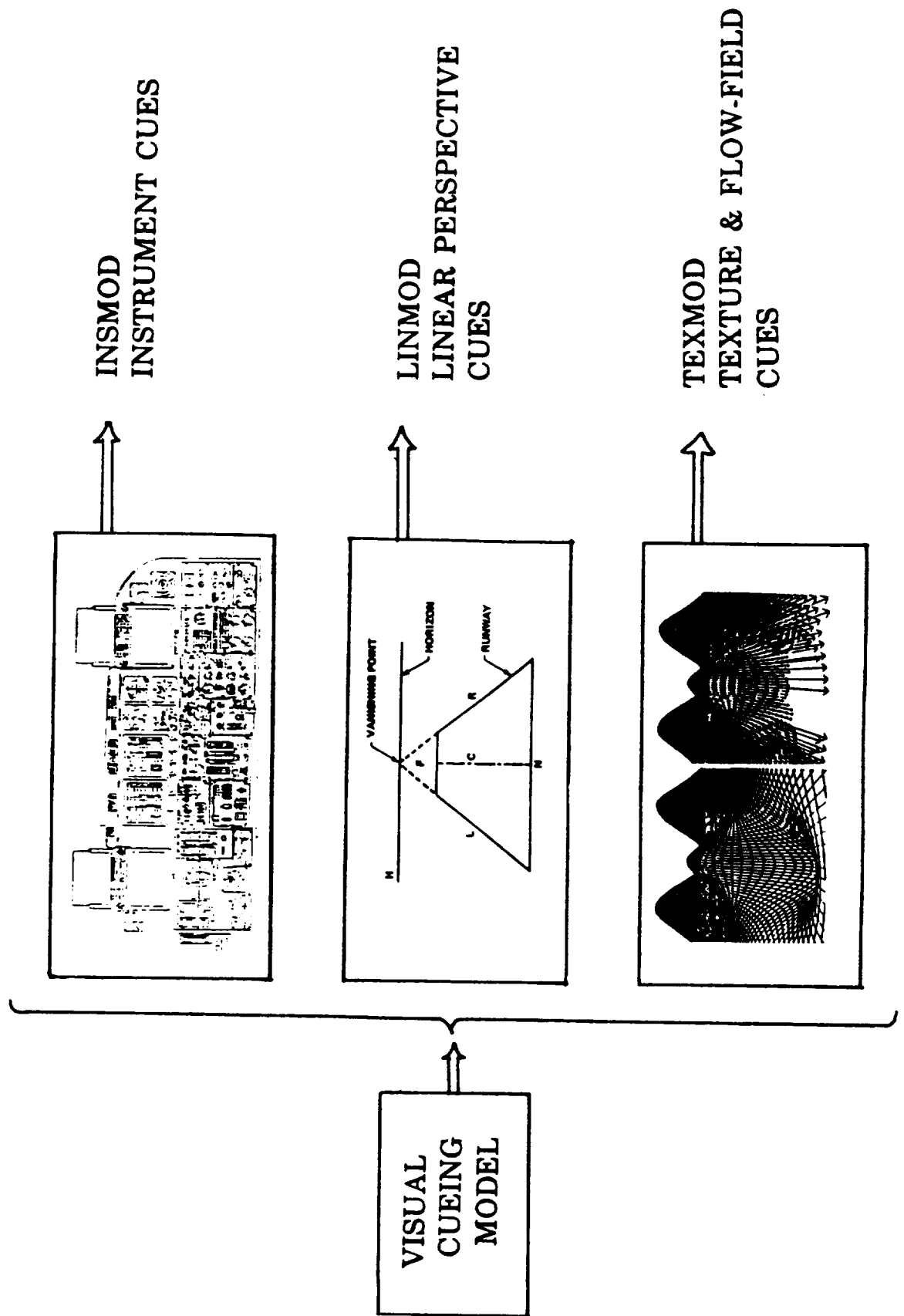
LINMOD: LINEAR PERSPECTIVE CUES

o **PILOT'S VIEW DURING LANDING APPROACH**
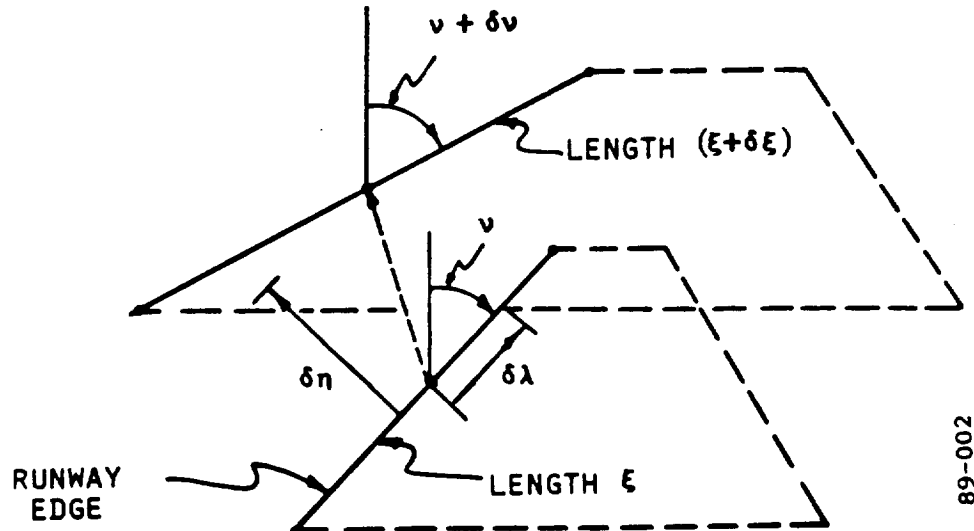
o **LINEAR PERSPECTIVE CUES**

- Length   :   scalar ξ , angular units
- Orientation: scalar ν , angular units wrt observer reference
- Location   :   vector (λ, η), angular units specifying
                 midpoint LOS

o **MODELING REQUIREMENTS**

- How does change in vehicle state (position/attitude) relate to
  change in cues?

- Find

$$\underline{Y}_{vis} \equiv (\xi, \nu, \lambda, \eta) = \underline{f}(\underline{x}) + \underline{v}_y$$

&

$$\delta \underline{Y}_{vis} = \frac{\partial \underline{f}}{\partial \underline{x}} \delta \underline{x} + \underline{v}_y$$

217

TEXMOD: TEXTURAL FLOW-FIELD CUES

o **PILOT'S VIEW DURING TF/TA**

static   dynamic
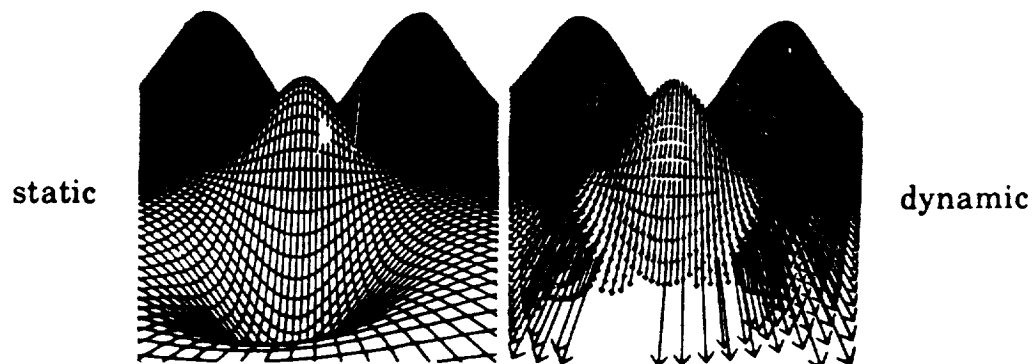
o **AIMPOINT AND SPIN AXIS ESTIMATION**

HORIZON   EXPANSION POINT   FLOW

ROTATION POINT   FLOW
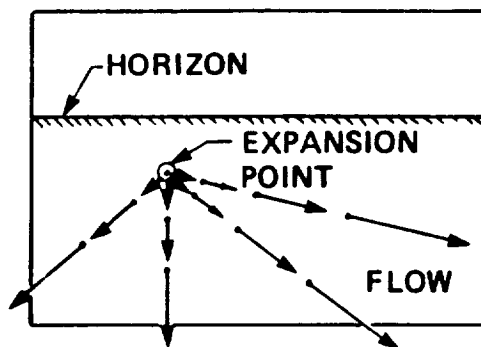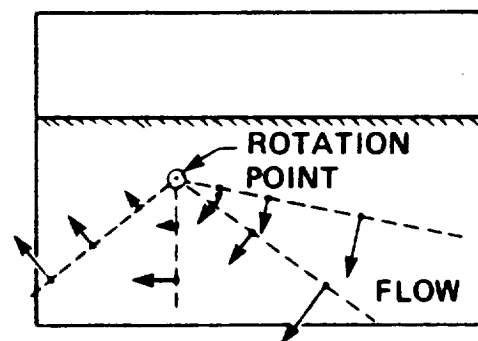
translation   rotation

o **MODEL OUTPUTS**

- Aimpoint
- Angular velocity
- Impact time map
- Relative orientation

$\hat{\omega}$ $\hat{u}$

T_1 T_2 T_3 P'_1 P'_2 P'_3

IMPACT TIME MAP
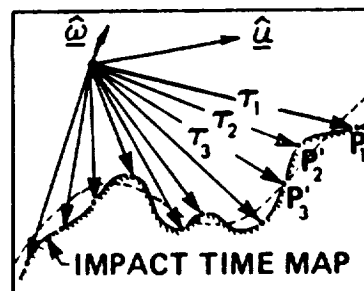
SIMPLE TERRAIN CUEING: TASK DESCRIPTION

o TASK: Altitude regulation against vertical gust

o DYNAMICS:

- Gust   : First Order Dryden, BW = 12 rad/s
- Vehicle: F-16 at SL, 400 kts, SAS-augmented

o DISPLAYS:
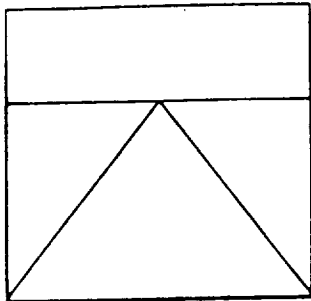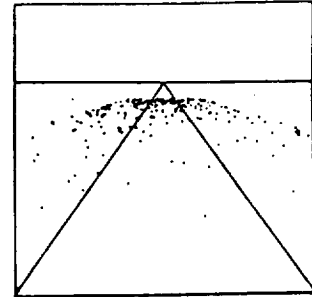
R: ROADWAY   T: TEXTURE   RT: ROADWAY &
                                               TEXTURE

o DISPLAY VARIABLES

- Roadway-only: $(\beta,\dot{\beta})$ from roadway
                     (θ,q) from horizon

- Texture-only: (h,γ) from textural flow
                     (θ,q) from pseudo-horizon

- Combined RT: $(\beta,\dot{\beta},h,\gamma,\theta,q)$

o VISUAL CUE THRESHOLDS

$(\beta,\dot{\beta})_{th}$ & $(\theta,q)_{th}$ from acuity estimates

$(h,\gamma)_{th}$ from textural flow model

o REFERENCE: WARREN & RICCIO (85); ZACHARIAS, WARREN
              & RICCIO (86)

219

SIMPLE TERRAIN CUEING: DATA & MODEL

o PERFORMANCE SCORES

Model

Data mean ± SD

(A,C) - Low Gain Display
(B,D) - High Gain Display
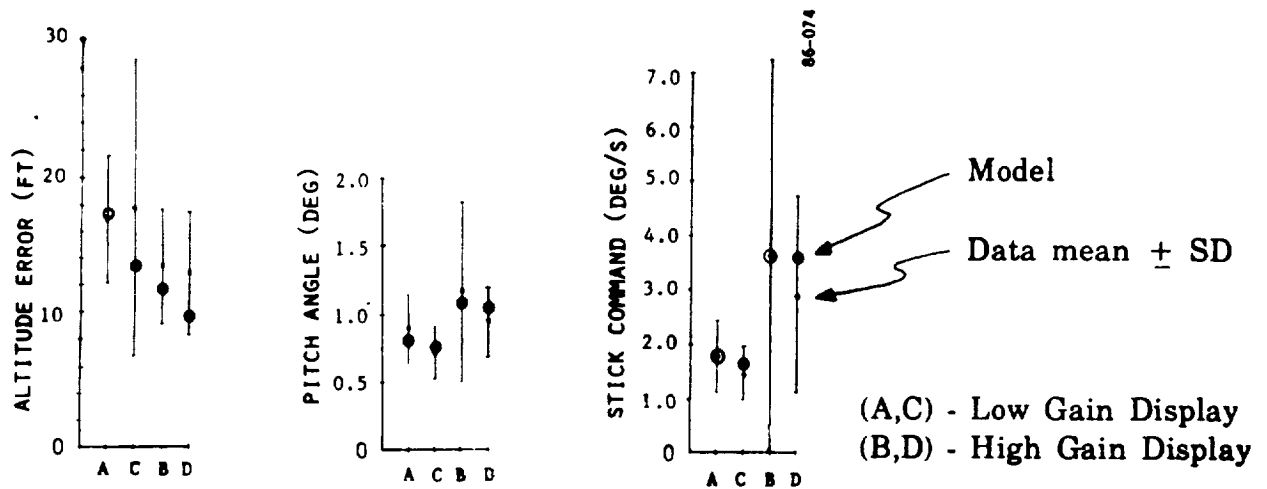
o PILOT FREQUENCY RESPONSE (stick/error)

FREQUENCY (RAD/SEC)
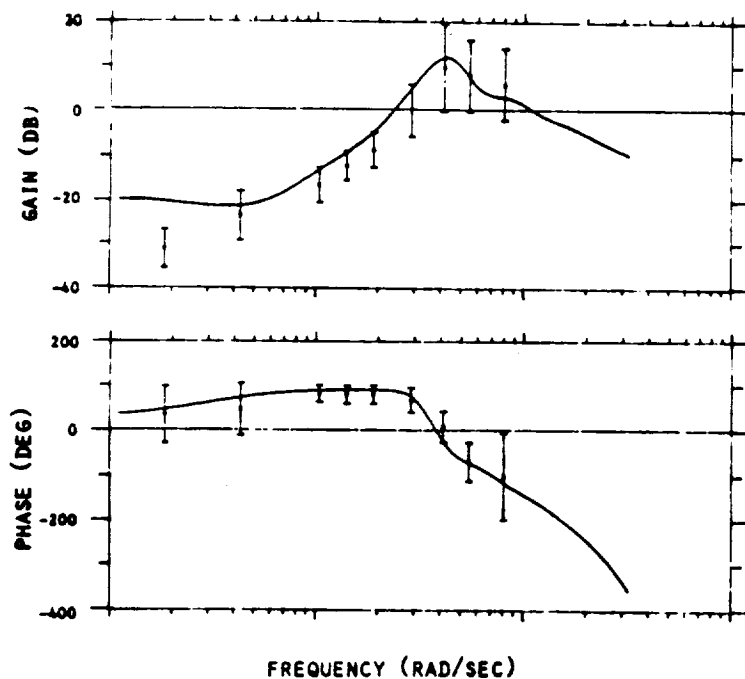
Condition B:
High Gain
Small Angle

220

PILOT MODEL PARAMETERS FROM DATA ANALYSIS

- DISPLAY VARIABLES

    - ROADWAY-ONLY: $(\beta_e, \dot{\beta}_e)$ FROM ROADWAY

      (θ, q) FROM HORIZON

    - TEXTURE-ONLY: (h_e, γ) FROM TEXTURAL FLOW

      (θ, q) FROM PSEUDO-HORIZON

    - COMBINED RT : $(\beta_e, \dot{\beta}_e, h_e, \gamma, \theta, q)$

- ATTENTION ALLOCATION

    70% ON HORIZON; 30% ON ROADWAY/TEXTURE

- VISUAL CUE THRESHOLDS

    $(\beta_e, \dot{\beta}_e)_{th}$ & $(\theta, q)_{th}$ FROM ACUITY ESTIMATES

    $(h_e, \gamma)_{th}$ FROM TEXMOD SIMULATIONS

- OBSERVATION NOISE RATIO: -18dB

- MOTOR PARAMETERS

    - TIME CONSTANT: $0.2s \rightarrow 0.4s$
    - MOTOR NOISE : $-40dB \longrightarrow -50dB$

- CENTRAL DELAY: 0.15s

PILOT MODEL PARAMETER VALUES FROM DATA ANALYSIS

PARAMETER	UNITS	DISPLAY TYPE ROADWAY (R)	TEXTURE (T)	COMBINED (RT)
MOTOR TIME CONSTANT τ_N				
L W GAIN (A,C)	SEC	0.30	0.40	0.30
HIGH GAIN (B,D)	SEC	0.20	0.35	0.20
MOTOR NOISE				
MOTOR NOISE, MN	DB	-50	-50	-50
PERCEIVED MOTOR NOISE, PMN	DB	-50	-40	-50
PROCESSING TIME DELAY τ_D	SEC	0.15	0.15	0.15
PERCEPTUAL NOISE LEVEL P_O	DB	-18	-18	-18
ATTENTION ALLOCATION				
HORIZON (θ, q)	--	0.70	0.70	0.70
ROADWAY $(\beta_e, \dot{\beta}_e)$	--	0.30		0.15
TEXTURE (h_e, γ)	--		0.30	0.15
VISUAL CUE THRESHOLDS				
HORIZON (θ_{th}, q_{th})	$(°, °/s)$	(1,.28)	(2,.56)	(1,.28)
ROADWAY $(\beta_{th}, \dot{\beta}_{th})$	$(°, °/s)$	(*,1)		(*,1)
TEXTURE (h_{th}, γ_{th})	$(FT, °)$		(**,2)	(**,2)

$$*\beta_{th} = (90° - \beta_a)/6 \qquad\qquad **h_{th} = 0.3h_a$$

SIMPLE TERRAIN CUEING: EXPERIMENTAL RESULTS & MODEL FINDINGS

o EFFECTS DUE TO DISPLAY TYPE:

  - Roadway-only provides adequate cues for task
  - Texture-only does also, but yields larger tracking errors, lower gains, greater lags, more remnant
  - Combined roadway-texture looks like roadway-only

o MODEL ANALYSIS FOLLOWED PERFORMANCE & FREQUENCY RESPONSE TRENDS ACROSS 3 DISPLAYS AND 4 FLIGHT CONDITIONS

  - Scores: Almost all within one SD
  - Gains/Phases: Almost all within one SD, but some gain mismatch at low frequencies
  - Remnant: Most within fraction of SD, but mid-frequency "plateau" missed

o DISPLAY EFFECTS MODELING

  - Roadway-only well-modeled by simple linear cue model
  - Texture-only modeled by TEXMOD-generated thresholds & increased motor time constant
  - Combined roadway-texture is dominated by roadway cues

223

SCENE GENERATOR DELAYS: TASK DESCRIPTION

o **TASK: FLY STRAIGHT & LEVEL AGAINST VERTICAL/LATERAL GUSTS**

o **OVERALL PILOT/VEHICLE BLOCK DIAGRAM**

89-006
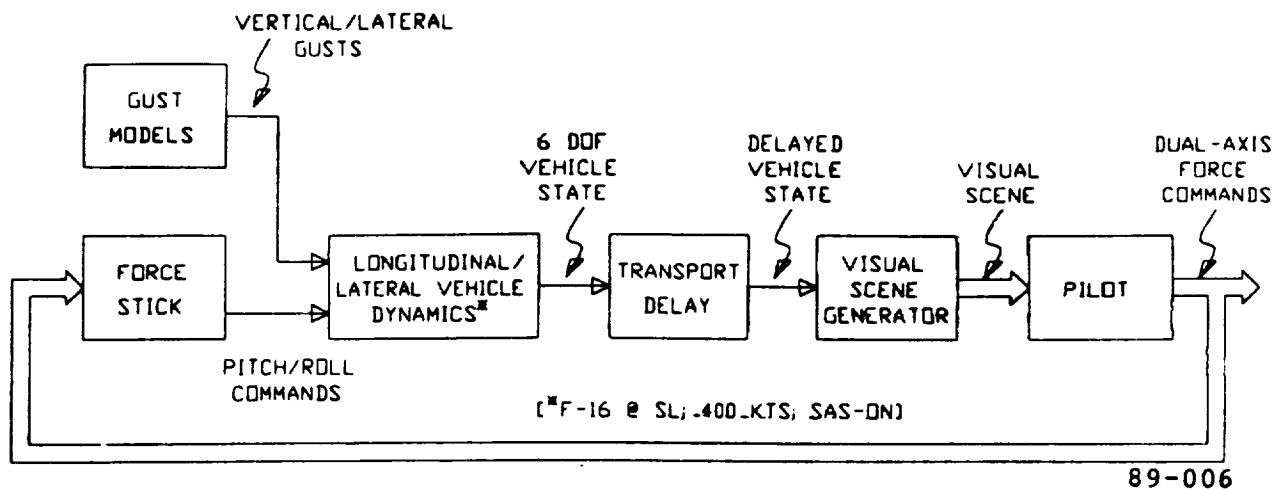
o **DELAY FACTORS: 50, 100, 200, 400 msec**

o **VISUAL SCENE**
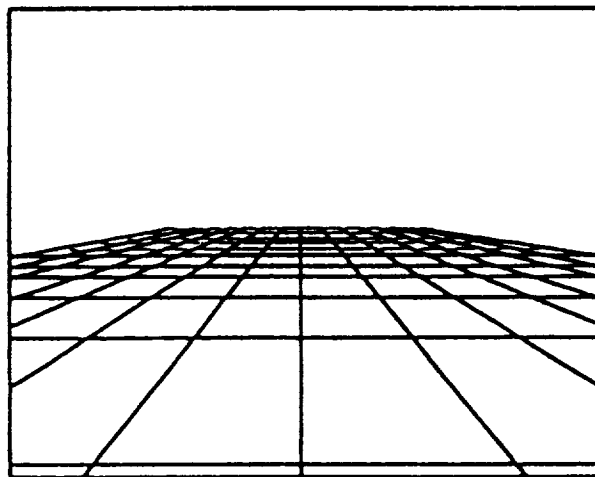
o **REFERENCE: RICCIO, CRESS, AND JOHNSON (87)**

224

SCENE GENERATOR DELAYS: MODEL ANALYSIS

o **TASK OBJECTIVE**

- Longitudinal subtask: minimize σ_h^2

- Lateral subtask: minimize $(\sigma_\psi^2 + k\sigma_Y^2)$

o **DYNAMICS MODEL**

- Linearized F16 6 DOF dynamics

- Sea level, 400 kts, SAS-on

o **DELAY MODEL**

Pade approximations to: 50, 100, 200, 400 msec delays

o **DISPLAY ANALYSIS**

- Meridian texture: (θ, h) & (ϕ, ψ, Y)

- Latitude texture: (θ, h) & (ϕ, ψ^*)

- Flow-field cues: rates of above

- Attention allocation set to optimize performance

- Thresholds set to zero

o **NON-DISPLAY PILOT PARAMETERS**

Fixed across conditions, except for increasing delay

DELAY EFFECTS ON PERFORMANCE: DATA & MODEL
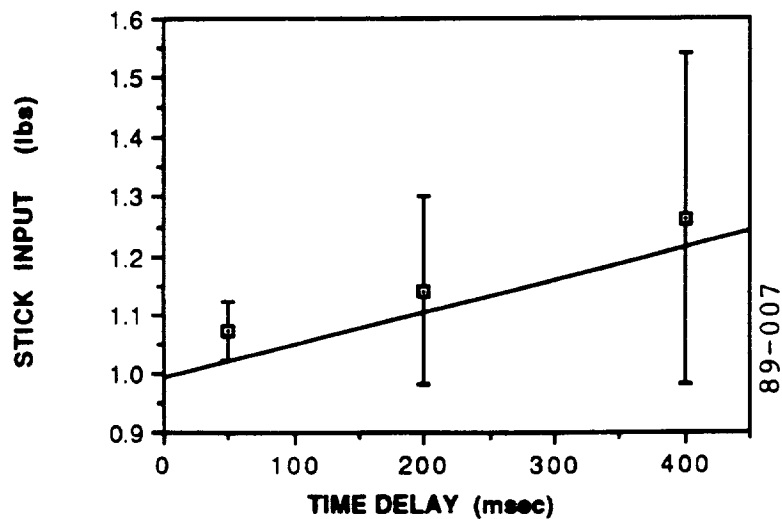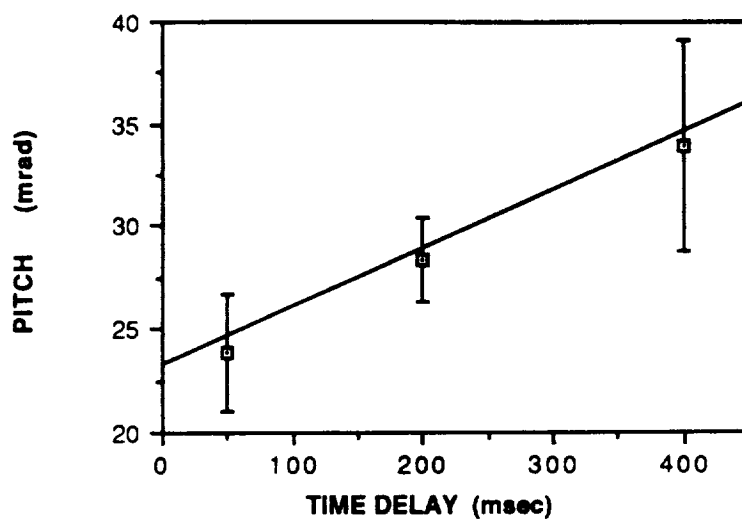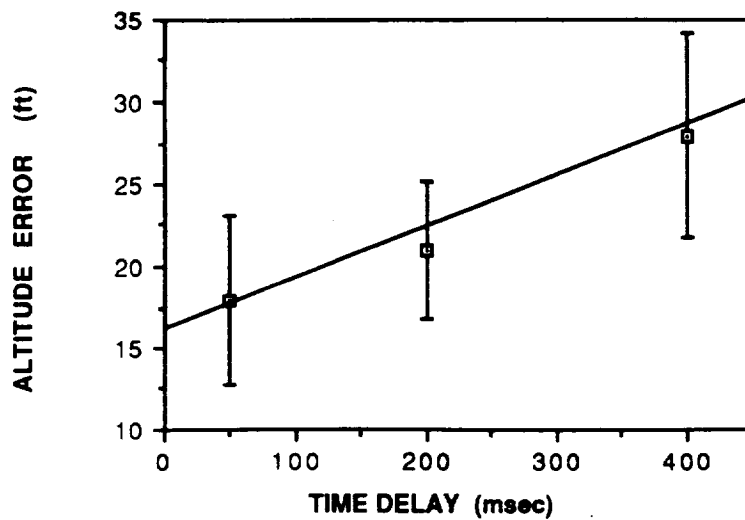
89-007

DELAY EFFECTS ON PILOT FREQUENCY RESPONSE: DATA & MODEL

o SHORT DELAY RESPONSE (50 msec)

o LONG DELAY RESPONSE (400 msec)
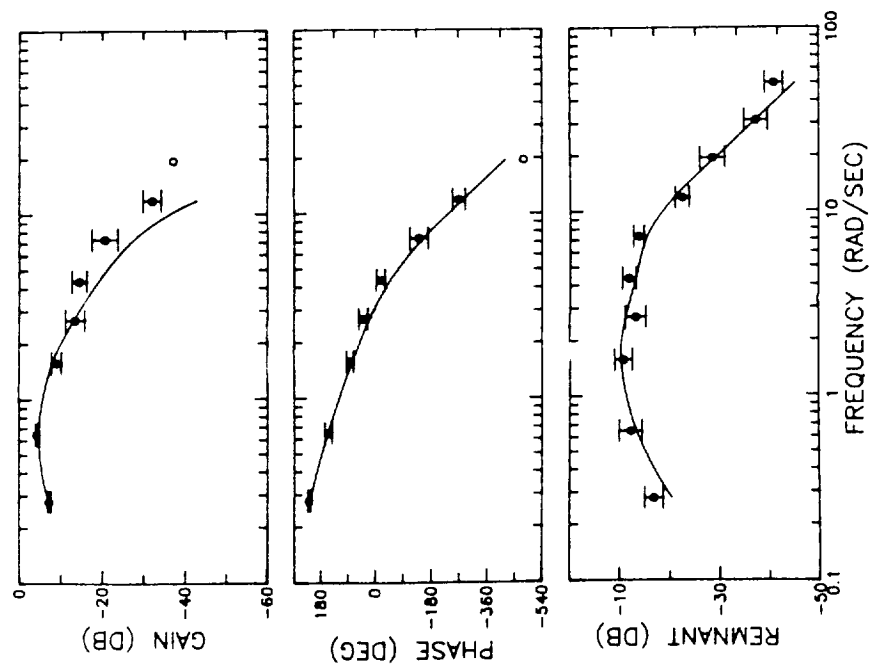
o stable gain
o increased lag
o increased remnant

227
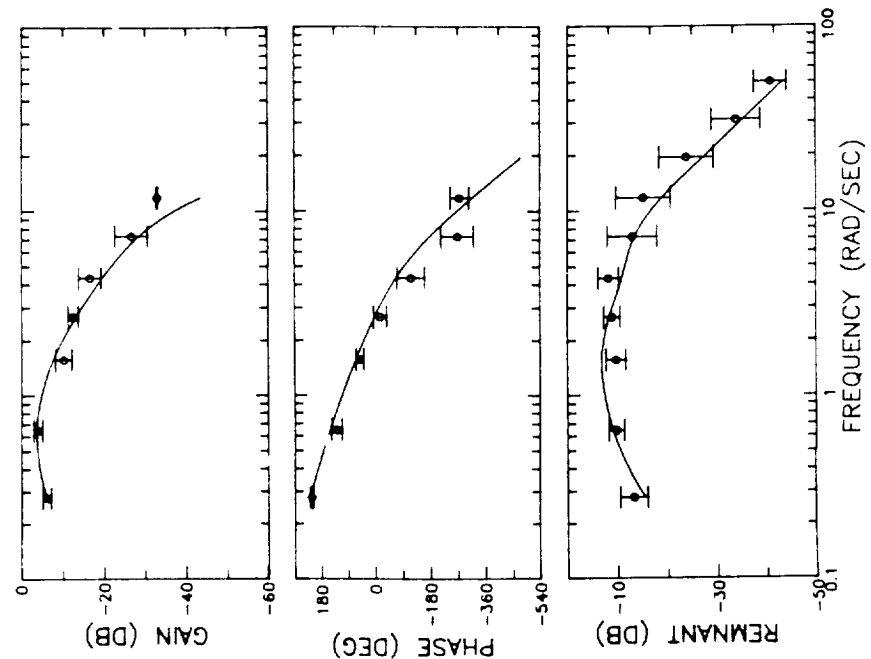
SCENE GENERATOR DELAYS: EXPERIMENTAL RESULTS & MODEL FINDINGS

o EFFECTS DUE TO INCREASING DELAYS

   - More stick activity & poorer performance

   - Increased response lags

   - Increased remnant (more random)

o MODEL ANALYSIS MATCH TO PERFORMANCE & FREQUENCY TRENDS

   - Performance trends with delays closely matched

   - Gain, phase, & remnant trends with frequency also well-matched

   - Obtained with fixed model parameters, except for increasing pilot delays

COCKPIT DISPLAY DESIGN: TASK DESCRIPTION

o TASK: LOW-LEVEL TERRAIN-FOLLOWING AT CONSTANT
HEADING

o DYNAMICS:

- Terrain: Second order matched terrain spectra

- Terrain-following guidance: Low order predictor

- Vehicle: B-1B at SL, Mach 0.85, SAS-augmented

o DISPLAY
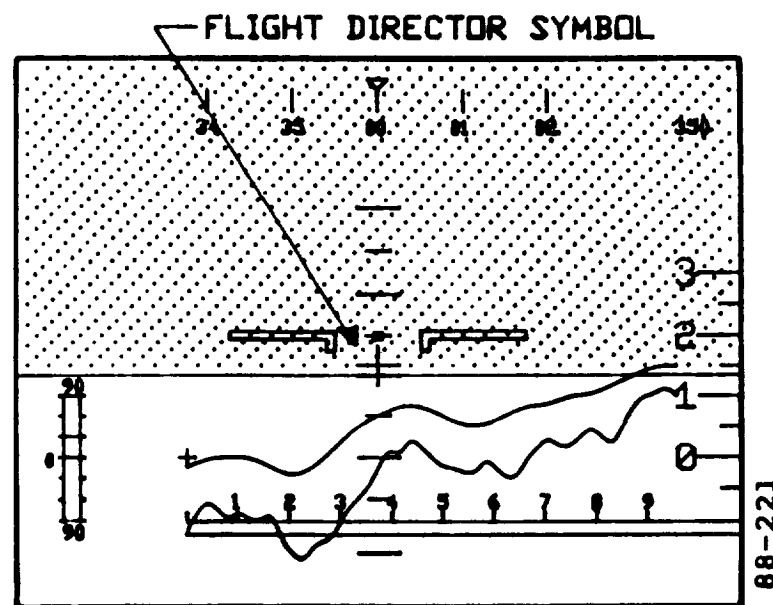
o DIRECTOR LAW

- Law: $\theta_{fd} = \alpha + \gamma_{dfp} - k * h_{error}$
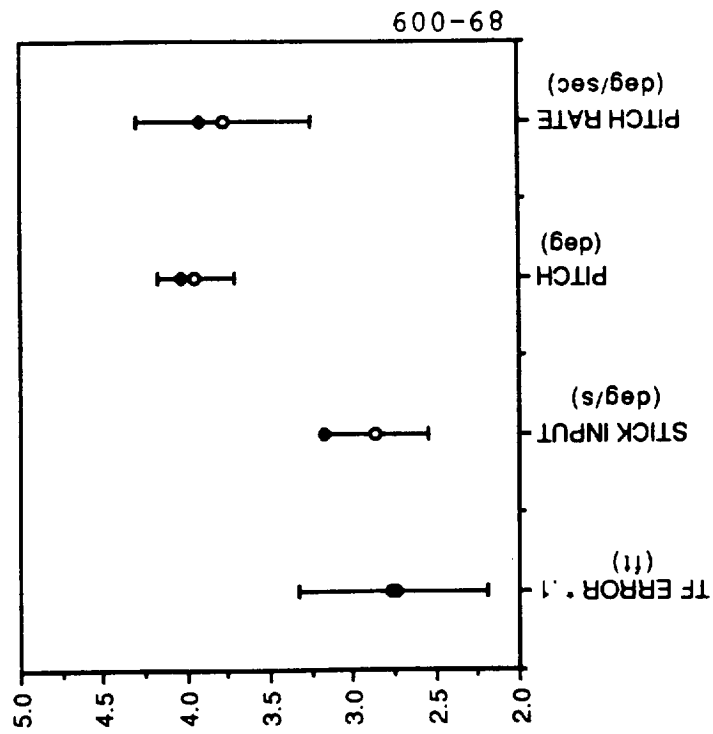
- Optimize director gain k

229

COCKPIT DISPLAY DESIGN: MODEL-BASED PROCEDURE

o CONDUCT PILOTED SIMULATION TO IDENTIFY BASELINE PILOT
  PARAMETERS

o SWEEP THRU DIRECTOR GAINS TO IDENTIFY OPTIMUM CHOICE

o CONFIRM CHOICE WITH SIMULATION USING OPTIMIZED
  DIRECTOR

o PRELIMINARY MODEL/DATA COMPARISONS (SINGLE SUBJECT)
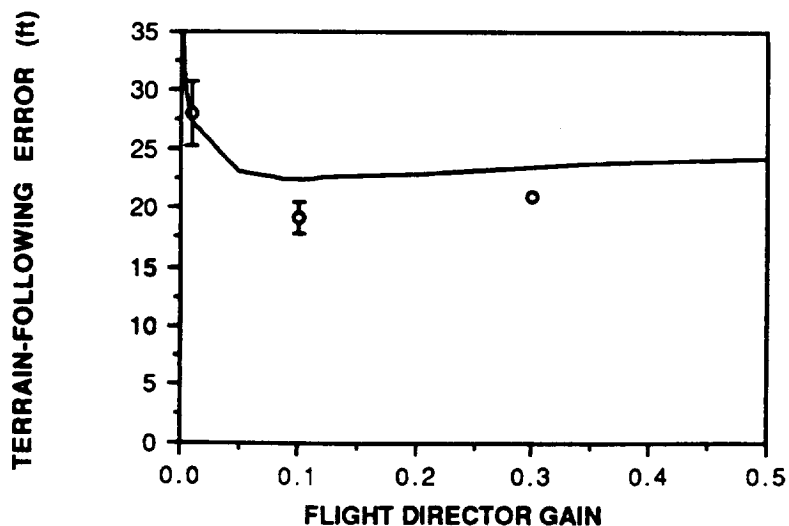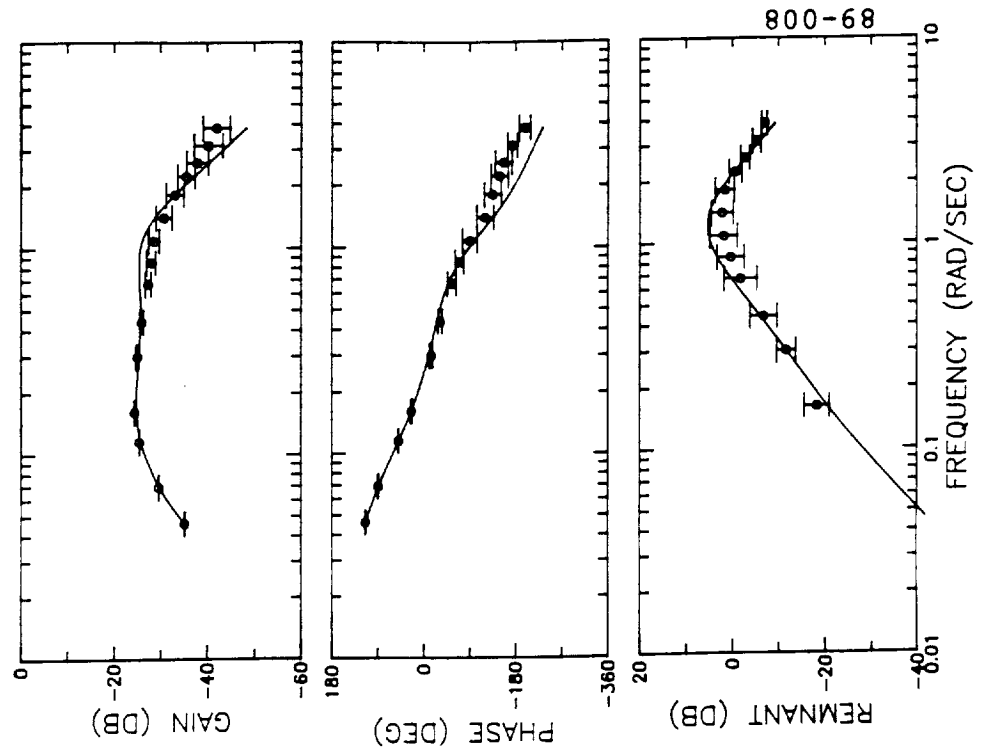
COCKPIT DISPLAY DESIGN: DATA AND MODEL

o   FREQUENCY RESPONSE
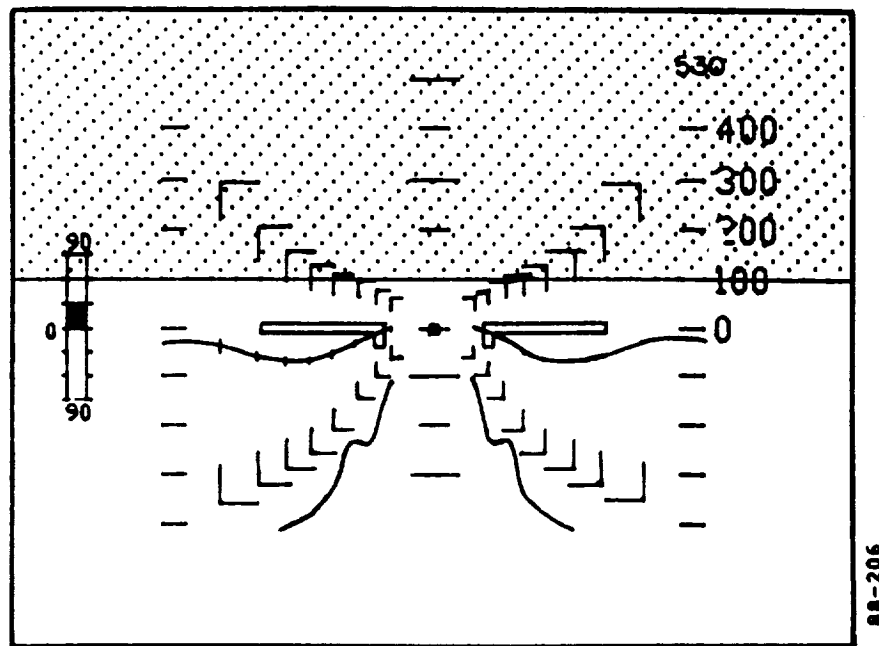

o   PERFORMANCE SCORES


231

BASELINE PICTORIAL GUIDANCE DISPLAY

o DISPLAY FORMAT

o KEY FEATURES

- Perspective view of TP & DFP overlaid on artificial horizon

- Artificial horizon gives attitude

- DFP-centered tunnel gives vertical/lateral path errors

- Tunnel dimensions indicate desired TF performance

- ADP gives high-gain TF error via indicator

- Path preview supports situational awareness

- Display integration minimizes attention-sharing

OPERATOR PERFORMANCE SCORES: VSD & PGD

233
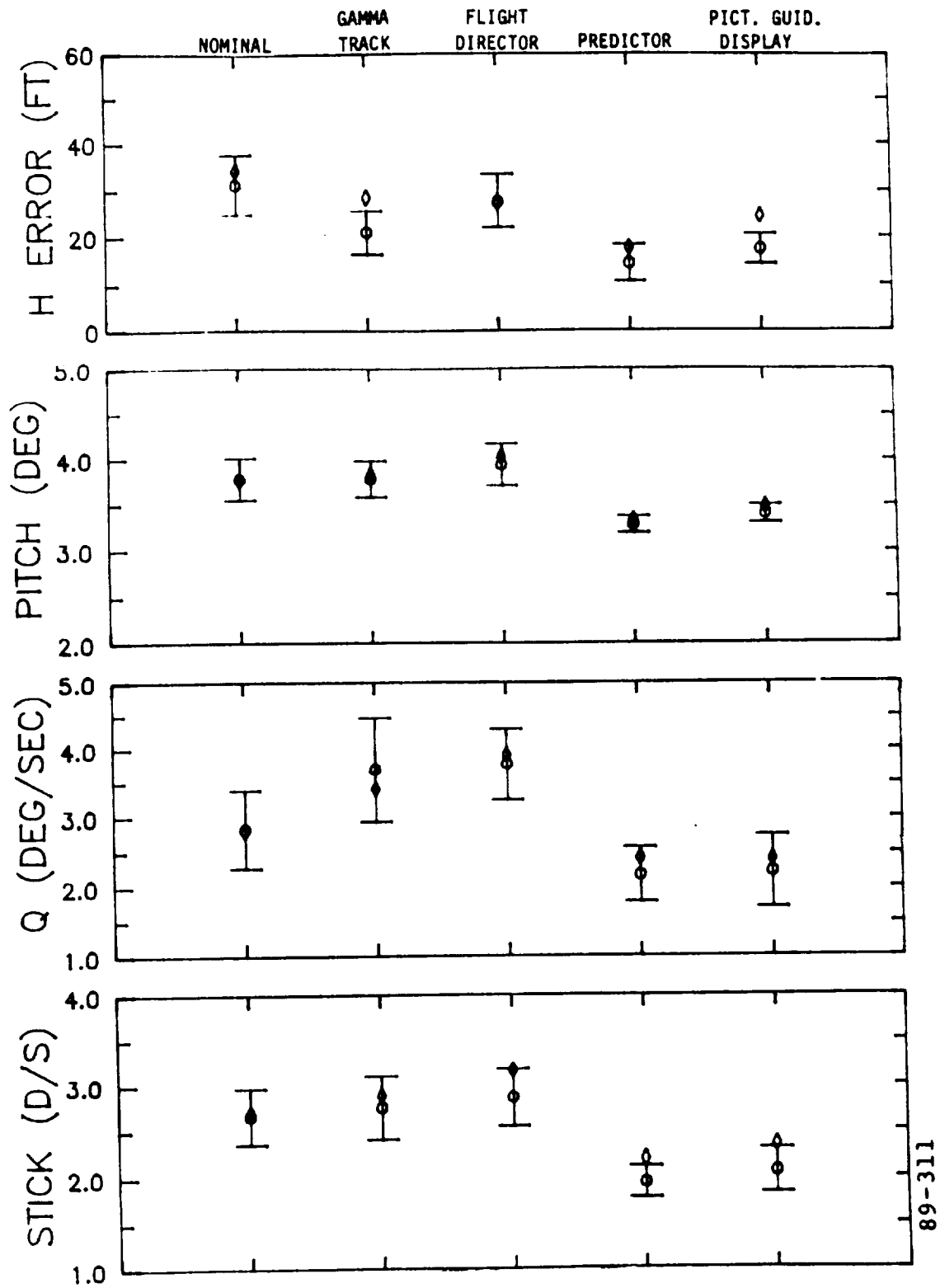
SUMMARY AND CONCLUSIONS

o SIMPLE TERRAIN CUEING DEMONSTRATES MODEL MATCH OF DOMINANCE EFFECTS

o SCENE GENERATOR DELAY TRENDS FOLLOWED VIA MODEL ANALYSIS

o MODEL-BASED DISPLAY DESIGN SUPPORTS DIRECTOR OPTIMIZATION

o GENERAL ROLE OF MODELING

- Provide structure and insight to multi-dimensional problem

- Provide means of data compression, interpolation, extrapolation

- Support design of focused (non-shotgun) experiments

- Support rational design of new displays

234

RANDOM THOUGHTS ON ROLE OF PILOT/VEHICLE MODELING

o DOES THE STRUCTURE GAINED BY MODELING OVERLY CONSTRAIN THE RESEARCH?

  - New experimental directions

  - New model development

o CAN EXCESSIVE COMPRESSION LEAD TO MISSED DATA TRENDS?

o ARE THE TECHNIQUES ADEQUATE TO ACCOUNT FOR OBSERVED BEHAVIOR? OR ARE THEY TOO LIMITED (e.g., LINEAR SYSTEMS)?

o DOES FUNCTIONAL EQUIVALENCE MISLEAD US REGARDING "TRUE" UNDERSTANDING OF THE PERCEPTION/CONTROL PROCESS?

APPENDIX B

LIST OF PARTICIPANTS

Dr. George J. Andersen
Department of Psychology
University of California
Riverside, CA 92507

Dr. James Cutting
Department of Psychology
Uris Hall
Cornell University
Ithaca, NY 14853-7601

Dr. John Flach
Department of Psychology
Wright State University
Dayton, OH 45435

Dr. Ronald Hess
Department of Mechanical, Aeronautical, and
  Materials Engineering
University of California
Davis, CA 95616

Dr. Larry Hettinger
Logicon Technical Services, Inc.
P.O. Box 317258
Dayton, OH 45431-7258

Dr. Ian Howard
Department of Psychology and Institute for Space
  and Terrestrial Science
York University
North York, Ontario M3J 1P3
Canada

Dr. Joe Lappin
134 Wesley Hall
Department of Psychology
Vanderbilt University
Nashville, TN 37240

Dr. Dean Owen
Department of Psychology
University of Canterbury
Christchurch 1, New Zealand

Dr. Dennis Proffitt
Department of Psychology
Gilmer Hall
University of Virginia
Charlottesville, VA 22903-2477

Dr. Gary Riccio
Department of Kinesiology
Rm 231 Freer Hall
University of Illinois
Urbana-Champaign, IL 61801

Dr. Rik Warren
Armstrong Aeromedical Research Laboratory
Human Engineering Facility
Wright Patterson AFB, OH 45433
(informal presentation only — no paper)

Dr. Lawrence Wolpert
Logicon Technical Services, Inc.
P.O. Box 317258
Dayton, OH 45431-7258

Dr. Greg Zacharias
Charles River Analytics, Inc.
55 Wheeler St.
Cambridge, MA 02138
(viewgraph presentation only — no paper.
  see Appendix)

NASA Ames Research Center
Aerospace Human Factors Research Division
Moffett Field, CA 94035-1000
    Vernol Battiste
    Dr. C. Thomas Bennett
    Dr. Stephen Ellis
    Sandra G. Hart
    Dr. Walter W. Johnson
    Dr. Mary Kaiser
    Dr. John Perrone

PRECEDING PAGE BLANK NOT FILMED

Report Documentation Page

1. Report No. NASA CP-3118	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Visually Guided Control of Movement		5. Report Date April 1991
		6. Performing Organization Code
7. Editor(s) Walter W. Johnson and Mary K. Kaiser		8. Performing Organization Report No. A-90200
		10. Work Unit No. 505-67-51
9. Performing Organization Name and Address Ames Research Center Moffett Field, CA 94035-1000		11. Contract or Grant No.
		13. Type of Report and Period Covered Conference Publication
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546-0001		14. Sponsoring Agency Code

15. Supplementary Notes

Point of Contact: Walter W. Johnson, Ames Research Center, MS 239-5
Moffett Field, CA 94035-1000
(415) 604-3667 or FTS 464-3667

16. Abstract

The papers in this volume were presented at an intensive, three-week workshop on visually guided control of movement. The participants were researchers from academia, industry, and government, with backgrounds in visual perception, control theory, and rotorcraft operations. The papers included invited lectures and preliminary reports of research initiated during the workshop. Three major topics are addressed: extraction of environmental structure from motion; perception and control of self motion; and spatial orientation. Each topic is considered from both theoretical and applied perspectives. Implications for control and display design are suggested.

17. Key Words (Suggested by Author(s)) Visual perception Vehicle control Self motion Spatial orientation	18. Distribution Statement Unclassified-Unlimited Subject Category – 53	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 246	22. Price A11